JMIR Serious Games

Contents

Original Papers

Viewpoints

Review

Serious Gaming Technology in Upper Extremity Rehabilitation: Scoping Review (e19071)

Original Paper

A Tablet App for Handwriting Skill Screening at the Preliteracy Stage: Instrument Validation Study

Linda Greta Dui[1], MSc; Francesca Lunardini[1], DPhil; Cristiano Termine[2], Prof Dr; Matteo Matteucci[1], Prof Dr; Natale Adolfo Stucchi[3], Prof Dr; Nunzio Alberto Borghese[4], Prof Dr; Simona Ferrante[1], Prof Dr

[1]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy

[2]Department of Medicine and Surgery, University of Insubria, Varese, Italy

[3]Department of Psychology, Università degli Studi di Milano-Bicocca, Milan, Italy

[4]Department of Computer Science, University of Milano, Milan, Italy

Corresponding Author:
Linda Greta Dui, MSc
Department of Electronics, Information and Bioengineering
Politecnico di Milano
Via Colombo 40
Milan, 20133
Italy
Phone: 39 0223999690
Email: lindagreta.dui@polimi.it

Abstract

Background: Difficulties in handwriting, such as dysgraphia, impact several aspects of a child's everyday life. Current methodologies for the detection of such difficulties in children have the following three main weaknesses: (1) they are prone to subjective evaluation; (2) they can be administered only when handwriting is mastered, thus delaying the diagnosis and the possible adoption of countermeasures; and (3) they are not always easily accessible to the entire community.

Objective: This work aims at developing a solution able to: (1) quantitatively measure handwriting features whose alteration is typically seen in children with dysgraphia; (2) enable their study in a preliteracy population; and (3) leverage a standard consumer technology to increase the accessibility of both early screening and longitudinal monitoring of handwriting difficulties.

Methods: We designed and developed a novel tablet-based app *Play Draw Write* to assess potential markers of dysgraphia through the quantification of the following three key handwriting laws: isochrony, homothety, and speed-accuracy tradeoff. To extend such an approach to a preliteracy age, the app includes the study of the laws in terms of both word writing and symbol drawing. The app was tested among healthy children with mastered handwriting (third graders) and those at a preliterate age (kindergartners).

Results: App testing in 15 primary school children confirmed that the three laws hold on the tablet surface when both writing words and drawing symbols. We found significant speed modulation according to size ($P<.001$), no relevant changes to fraction time for 67 out of 70 comparisons, and significant regression between movement time and index of difficulty for 44 out of 45 comparisons ($P<.05$, $R^2>0.28$, 12 degrees of freedom). Importantly, the three laws were verified on symbols among 19 kindergartners. Results from the speed-accuracy exercise showed a significant evolution with age of the global movement time (circle: $P=.003$, square: $P<.001$, word: $P=.001$), the goodness of fit of the regression between movement time and accuracy constraints (square: $P<.001$, circle: $P=.02$), and the index of performance (square: $P<.001$). Our findings show that homothety, isochrony, and speed-accuracy tradeoff principles are present in children even before handwriting acquisition; however, some handwriting-related skills are partially refined with age.

Conclusions: The designed app represents a promising solution for the screening of handwriting difficulties, since it allows (1) anticipation of the detection of alteration of handwriting principles at a preliteracy age and (2) provision of broader access to the monitoring of handwriting principles. Such a solution potentially enables the selective strengthening of lacking abilities before they exacerbate and affect the child's whole life.

XSL•FO

RenderX

Introduction

Background

Dysgraphia is a learning disability that involves unsatisfactory handwriting production, given age, normal intelligence, and absence of neurological, perceptual, or motor problems [1]. Dysgraphia can also be characterized by an excessively long execution time notwithstanding adequate readability, resulting in difficulty to keep up with peers in everyday tasks or homework [2,3]. Dysgraphia heavily impacts a child's school and everyday life, as handwriting is fundamental in the learning process. Handwriting difficulties, especially at an early educational age, have an effect on children's self-esteem and cause behavioral problems and early school abandonment [4]. Dysgraphia can occur in isolation or in association with a specific learning disorder (eg, developmental dyslexia with a moderate comorbidity correlation) [1] or fall within a more generalized developmental coordination disorder [5]. In the latter case, motor coordination difficulties are not limited to writing, but emerge in other daily activities (eg, playing an instrument, performing sports activities, and lacing up shoes).

The exact epidemiology of dysgraphia is still under debate, resulting in differences in the diagnostic criteria adopted by different schools of thought [6]. Beside the lack of consensus, current diagnostic tools are affected by the following two main limitations: (1) as they are primarily based on handwriting production, the actual diagnosis cannot be performed before complete handwriting maturation in the third year of primary school [7] and (2) the examination itself is based on either qualitative observation of the writing outcome, which might be prone to different interpretations according to the examiner, or the writing speed, which is not a comprehensive indicator.

The limitations of the diagnostic procedure contribute to uncertainty in the estimation of the actual incidence of the phenomenon. Indeed, the reported prevalence of handwriting problems ranges between 5% and 27% [6,8]. Such a wide range includes both underestimation, when a child has no access to a diagnostic visit with a specialist, and overestimation, when the lack of a prompt intervention exacerbates transient difficulties, which are improperly translated into the corresponding disability.

Needs

Against this background, there is an urgent need for innovative solutions to support the screening for dysgraphia. An effective solution must meet the following requirements: (1) quantitatively reveal potential alterations in writing production; (2) detect potential weaknesses at an early stage, thus enabling early intervention; and (3) be easily accessible to the user in order to reduce delayed diagnoses and allow longitudinal monitoring.

To meet the first requirement, understanding in which way dysgraphia alters the principles of handwriting is fundamental to quantify such alterations. It is known that handwriting problems are often linked to a lacking sense of rhythm [9,10].

In handwriting, rhythm can be translated into the ability of pacing letters to keep constancy in movement execution. Specifically, the following two laws of motion relevant to this scope have been studied during handwriting: *isochrony* and *homothety*. The *isochrony* principle states that bigger gesture execution is accompanied by an increase in average movement speed to keep the movement time approximately constant [11]. The *homothety* principle predicts that the fraction of time devoted to each letter of a word is kept constant and is independent of the total word duration [11,12]. When both these principles are fulfilled, relative and total execution time are approximately constant between different word sizes or speeds. Both these principles are reported to be altered in handwriting production among pupils with dysgraphia having dyslexia comorbidity [13]. Indeed, when asked to write a word on a piece of paper affixed on a digitizing tablet at different speeds and dimensions, primary school children affected by learning disabilities were not able to adapt their speed to task requirements, showing difference from their typically developed peers.

Experts stress the concept that pupils with dysgraphia tend to write every single letter faster than typically developing peers, with greater pauses and poorer accuracy [2]. Such observations suggest an alteration of the *speed-accuracy tradeoff* principle, which states that the more accurate the task, the longer it takes to accomplish it. Smits-Engelsman et al [14] reported that children with learning disabilities were less accurate in continuous cyclical fast aiming tasks, whereas they performed similarly to their peers when paced by an external "go" signal. The speed-accuracy tradeoff alteration in this group seems therefore related to continuous open loop activities, such as handwriting tasks. A model used to assess the speed-accuracy tradeoff during continuous movements is the *steering law* [15], a continuous formulation of the Fitts' law [16] with a shift of the paradigm from pointing to path steering. Such a law was tested with different geometrical constraints, but was never applied to handwriting.

To meet the second requirement, uncovering the laws underlying handwriting in drawing is crucial. This would allow the detection of potential difficulties before the acquisition or at least before the complete maturation of the writing gesture. Previous studies on isochrony [17,18] confirmed that such a principle is characteristic of gesture production, thus allowing its study in drawings. As for homothety, an attempt of proving its presence before the complete maturation of handwriting was made by Pagliarini et al [19]. They investigated homothety during handwriting in children from the first to the fifth grade, revealing that the principle holds overall, but a deviation was observed for younger children. This finding leaves room for a further investigation on homothety at a preliteracy age.

Finally, to meet the third requirement, leveraging standard consumer technology is key to allowing easy access to the screening process, as reported in other fields [20]. In the scientific literature, the classical setup for quantitative and

reliable handwriting evaluation involves digitizing tablets, which are still far from being deployed in nonclinical environments. Indeed, they are more expensive and less general purpose than standard consumer technologies, such as tablets. Moreover, data recording and fruition require specialized software, preventing the acquisition protocol from being straightforward to nonexpert users. For these reasons, tablet-based serious games have been recently proposed to both test perceptual motor skills in primary school children with and without suspected learning difficulties [21,22] and propose preliminarily treatments for dysgraphia [23]. Indeed, the use of tablet-based solutions entails many advantages like the opportunity of introducing encouraging feedback on the user's performance and the possibility of accomplishing increased interaction and more dynamic scenarios. Such features contribute to achieving a greater engagement, which is key especially for children's application. However, such a technology must be validated for the study of handwriting, since the tablet surface can affect writing performance. Difficulties in changing the writing support are particularly evident in children [24], and we must assure that a change in the writing support does not alter the handwriting laws that hold for the pen-and-paper standard approach.

Objectives

Given the highlighted needs, we devised and developed *Play Draw Write*, a tablet app to investigate handwriting through serious games designed to test isochrony, homothety, and speed-accuracy tradeoff principles both in word writing and symbol drawing. We proposed the following three goals: (1) to investigate if these writing principles hold when writing on a tablet surface; (2) to verify that symbol drawing can anticipate the detection of potential difficulties at the preliteracy stage; and (3) to investigate how writing-related skills and principles change with gesture acquisition.

Methods

Materials

The hardware chosen for the experiment was Samsung Galaxy Tab A with an S-Pen (Figure 1). The S-Pen with a rubber tip was used to better mimic the paper-and-pencil condition.

Figure 1. Experimental setup. A kindergartner is executing one of the exercises on the tablet with an S-Pen.


The app was developed in Unity 2018.3.2f1 following the principles of theory-driven evidence-based serious games for health [25,26]. We set the sampling rate at 50 Hz, as the frequency content of the recorded movement was expected to be less than 22 Hz [27].

The app *Play Draw Write* presented the following two exercises: *copy game* and *tunnel game*, which can be seen in Multimedia Appendix 1 and Multimedia Appendix 2, respectively. The copy game was aimed at testing both isochrony and homothety. We presented the participants an interface with an empty canvas and an example of the word and symbols they had to copy, together with the execution modality (Figure 2A). At the end

of each execution, a new empty canvas appeared on the screen. As for the word, "mele" ("apples" in Italian) was chosen since it contains the letters E and L, which are useful to test praxical abilities [28]. As for the symbols, following the Denver test [29], a circle, square, and triangle were chosen. In addition, a sequence of these symbols was also used to test for homothety in symbols.

Figure 2. App interface. A: copy game; B: tunnel game.
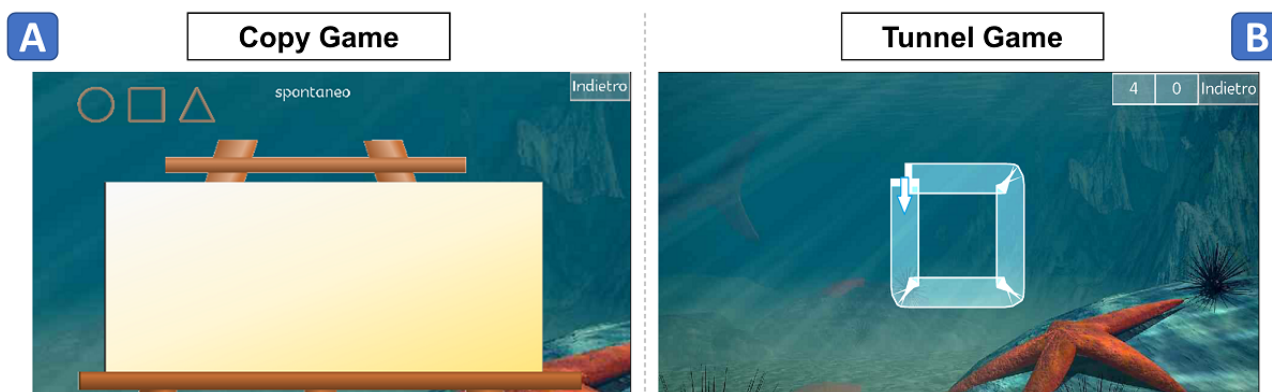
The tunnel game was aimed at testing the speed-accuracy tradeoff for both words and symbols. Our purpose was to detect the adaptability to different task requirements (ie, accuracy). According to the steering law, accuracy is defined by the index of difficulty (ID) as follows:

$$ID = A / W \ (1)$$

where W is the width between the borders of the path and A is the amplitude of the gesture (ie, perimeter for symbols and the path length for words, both measured in the center of the path). We combined A and W to obtain five different IDs for words and five for symbols, following clinicians' guidelines (Multimedia Appendix 3). We presented the participants an interface with symbol- and word-shaped paths with different IDs (Figure 2B). An arrow indicated the direction to follow for symbol steering. If the border of the tunnel was crossed, a noise alerted participants to correct themselves. At the end of each execution, a new word/symbol appeared on the screen, with different accuracy constraints. Both for words and symbols, tutorials were provided to familiarize with the task. In the tutorial, a cartoon hand reproduced the exercise. Regarding words, paths were shaped as the word "ele." Regarding symbols, paths were shaped as circles and squares. To prevent fatigue, the triangle was discarded from the tunnel game, as it is known to be the last symbol learned by children [29].

Participants and Protocol

Data collection was performed on children attending the third grade of a primary school and the last year of a kindergarten in the province of Como, Italy. Target ages were selected as handwriting is mastered in third grade [7] and not yet learned in kindergarten. The inclusion criteria were as follows: both right and left handedness, and normal graphical abilities for kindergarten children, according to teachers' judgement. The exclusion criteria were as follows: known cognitive or

sensory-motor problems and learning disability diagnosed for primary school children. All children who attended the partner schools and met the inclusion criteria were invited to participate in the study.

A minimum sample size of 10 children per group was computed on the basis of previous experiments in the field of the speed-accuracy tradeoff [30]. We considered 80% power, a first-type error of 5%, and a drop-out rate of 20%.

We collected data at the end of the school year. Acquisition took place in a quiet room with the child comfortably seated. The complete protocol (Figure 3) duration was 15 to 30 minutes, according to the ability of the children.

In the copy game, the instructions were to copy the word/symbol shown in the upper part of the screen, in the indicated modality (Figure 2A), as in previous studies on similar tasks [13]. We also orally specified the modality, as kindergartners cannot read yet. Word copying was executed by primary school children only, as kindergartners cannot write yet. In this exercise, we specified that they could use their own handwriting style, as trying to copy in an unnatural handwriting would have biased the results. In the tunnel game, during the tutorial, children were instructed to steer the path twice for symbols and once for words, as fast as they could, without crossing the external border. If they accidentally drew outside the path, they were encouraged to correct themselves, without stopping the execution. The instruction gave equal importance to speed and accuracy, as per the theory of constant throughput in speed-accuracy tradeoff [31].

The experimental procedure was approved by the Ethics Committee of Politecnico di Milano (n.10/2019). Written consent was received from school deans, parents, and primary school children.

Figure 3. Study protocol. A: copy game; B: tunnel game.
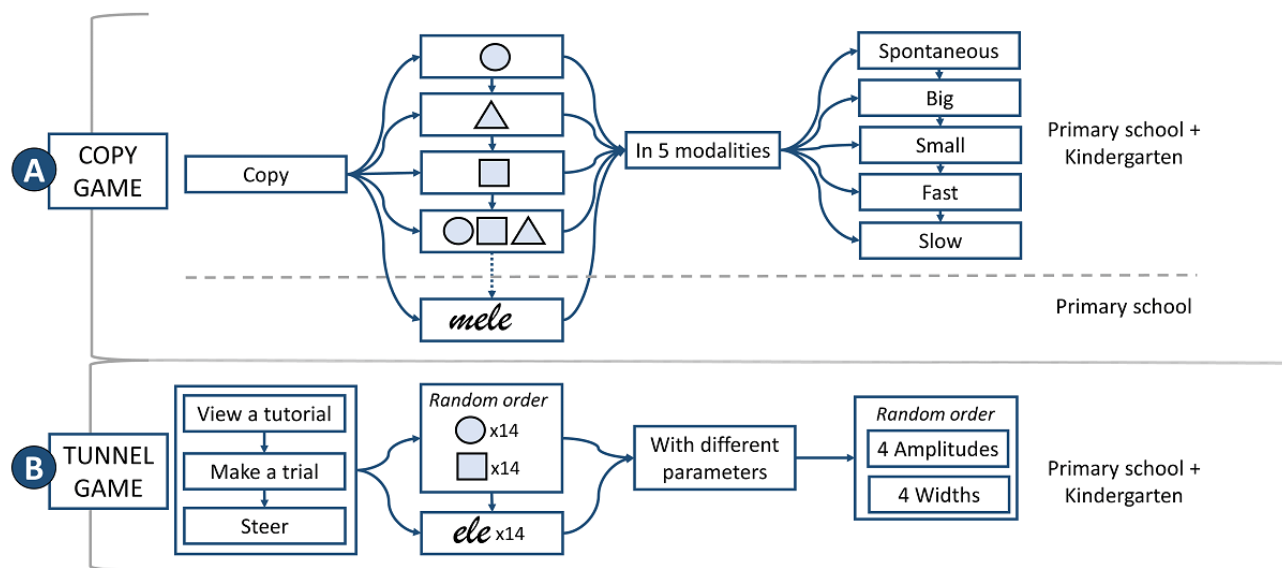
Data Analysis and Statistics

Data analysis was performed in Matlab R2018b (MathWorks) and R 3.3.3 (R Development Core Team). Significance was set at .05. Given the small sample size, nonparametric statistics were selected for all tests.

Copy Game

First, instruction compliance was tested using a Wilcoxon matched paired test comparing the trajectory length between the *big* and *small* modalities.

To test isochrony, separately for words and symbols, we considered between-condition differences in the average execution speed. To be compliant with the isochrony principle, we expected the *big* and *fast* modalities to be executed faster than the *small* and *slow* modalities. We first verified the validity of the law in primary school children (words and symbols) and then extended it to kindergartners to test the potential influence of gesture acquisition.

We considered speed, rather than time, to normalize the execution time for the size, as it was self-selected. For both words and symbols, we discarded the first and last 5% of the trajectory to avoid border effects. Thereafter, we computed speed as the ratio between the discrete difference between adjacent coordinates and the corresponding timestamps. We filtered the result with a 10-Hz low-pass filter, and we averaged the speed over the entire path to obtain a single value per execution. To test the effect of the modality on speed, we used Friedman and Bonferroni *post-hoc* tests.

To understand if the primary school and kindergarten groups approached the exercise differently, we performed, separately

for each symbol, a Mann-Whitney *U* test with the *spontaneous* modality speed as the dependent variable and the age group as the independent variable.

To test for homothety on both word and symbol sequences, we computed the percentage of time (*fraction time*) dedicated to each element of the sequence. For cursive words, the lowest point between two letters was considered as the transition point. Concerning symbols, the first and last frames the pen was on the screen were considered as the beginning and the end of a symbol, respectively. A Friedman test was performed separately on words and symbols (independent variable: modality; dependent variable: fraction time). In case of significance, Bonferroni *post hoc* was executed. No effect of the modality was expected.

Tunnel Game

To study the speed-accuracy tradeoff during handwriting, the steering law was adopted. The steering law predicts a linear relationship between movement time (MT) and ID as follows:

$$MT = a + b\ ID\ (2)$$

To compute MT in the word "ele," we considered the timestamp difference between the first "*E*" crossing and the last "*E*" crossing (Figure 4).

To compute MT in symbols, we considered the central part of the execution as follows: for the circle, we discarded the first and last semicircles, leaving a complete central circle; for the square, we discarded the first two and the last two sides, leaving a complete central square. We discarded trials where more than 40% of the path was outside the borders [31], in respect to the total length of the trace (Figure 4; error rate: 42.76%).

Figure 4. Speed-accuracy tradeoff exercise. The transparent background represents the app interface.
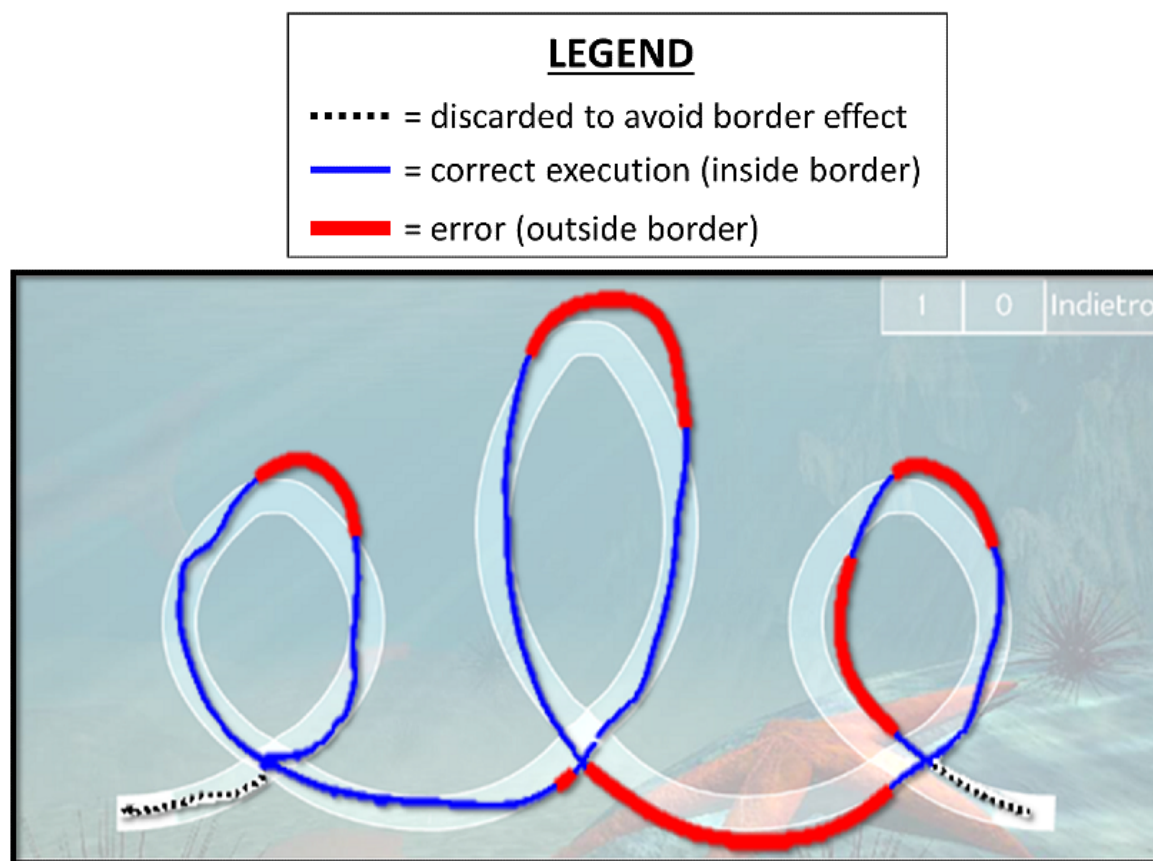
Separately for the word and the two symbols, we calculated the *global MT*, that is, the median movement time across all IDs, for each subject. For the word and the two symbols, we computed a linear regression between IDs, as an independent variable, and MT, as a dependent variable, for each subject separately. For each regression, we computed the R^2 and we checked for regression significance that implies the validity of the law. Moreover, we evaluated significant fittings using the root mean square error (RMSE) of regression as an index of the goodness of fit. Significant fittings only were used to compute the inverse of the regression line slope (the index of performance [IP]), which was computed as follows:

$$IP = 1 / b \ (3)$$

The IP is an index of adaptation to a task's accuracy requirements [16]. We tested for differences in MT, RMSE, and IP, considering both age group (Mann-Whitney test) and symbols (Friedman test) as independent variables. When significance was reached, a Bonferroni *post-hoc* test was performed.

Results

Participants

We enrolled 15 right-handed primary school children (7 males and 8 females; mean age 8.8 years, SD 0.3) and 19 kindergartners (9 males and 10 females; 2 left handed; mean age 5.9 years, SD 0.4).

Data Analysis and Statistics

Test results are reported in Multimedia Appendix 4 and Multimedia Appendix 5, and in Figures 5-7.

Figure 5. Isochrony. A: word, primary school; B: symbols, primary school; C: symbols, kindergarten. X-axis: execution modality; Y-axis: mean speed as box-and-whiskers plots. The red horizontal line is the median, the notch is its 95% CI, and the box is the IQR. Asterisks for between-condition differences involving the spontaneous modality are not reported for clarity purposes.*P<.05, **P<.005, ***P<.001.
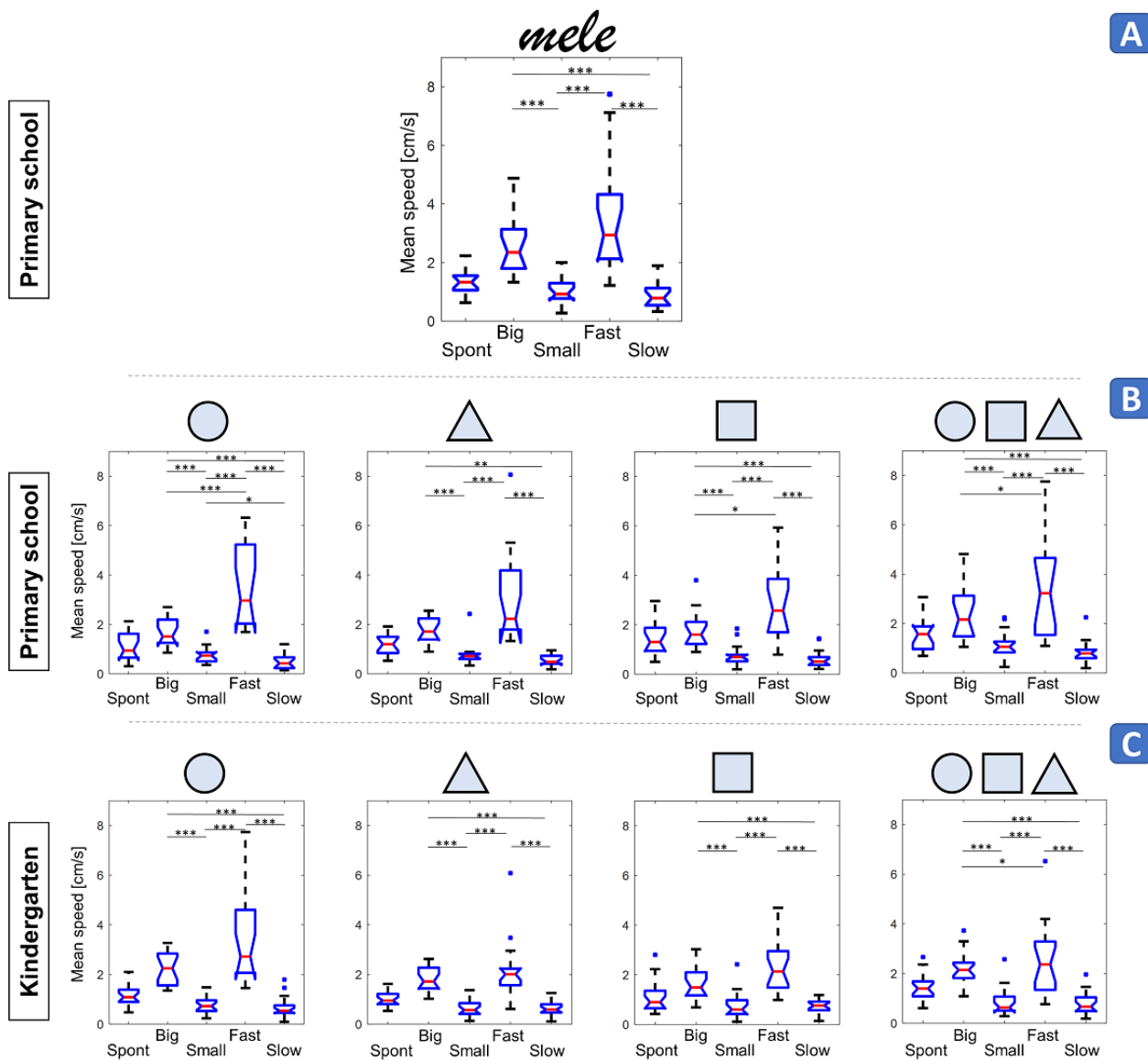
Figure 6. Homothety. A and B: primary school; C: kindergarten. A: word; B and C: symbols. Each subplot represents one element of the sequence of letters or symbols. X-axis: execution modality; Y-axis: fraction time, in normalized unit (1 represents the total execution time) as box-and-whiskers plots. The red horizontal line is the median, the notch is its 95% CI, and the boxes represent the IQR. *$P<.05$, **$P<.005$, ***$P<.001$.
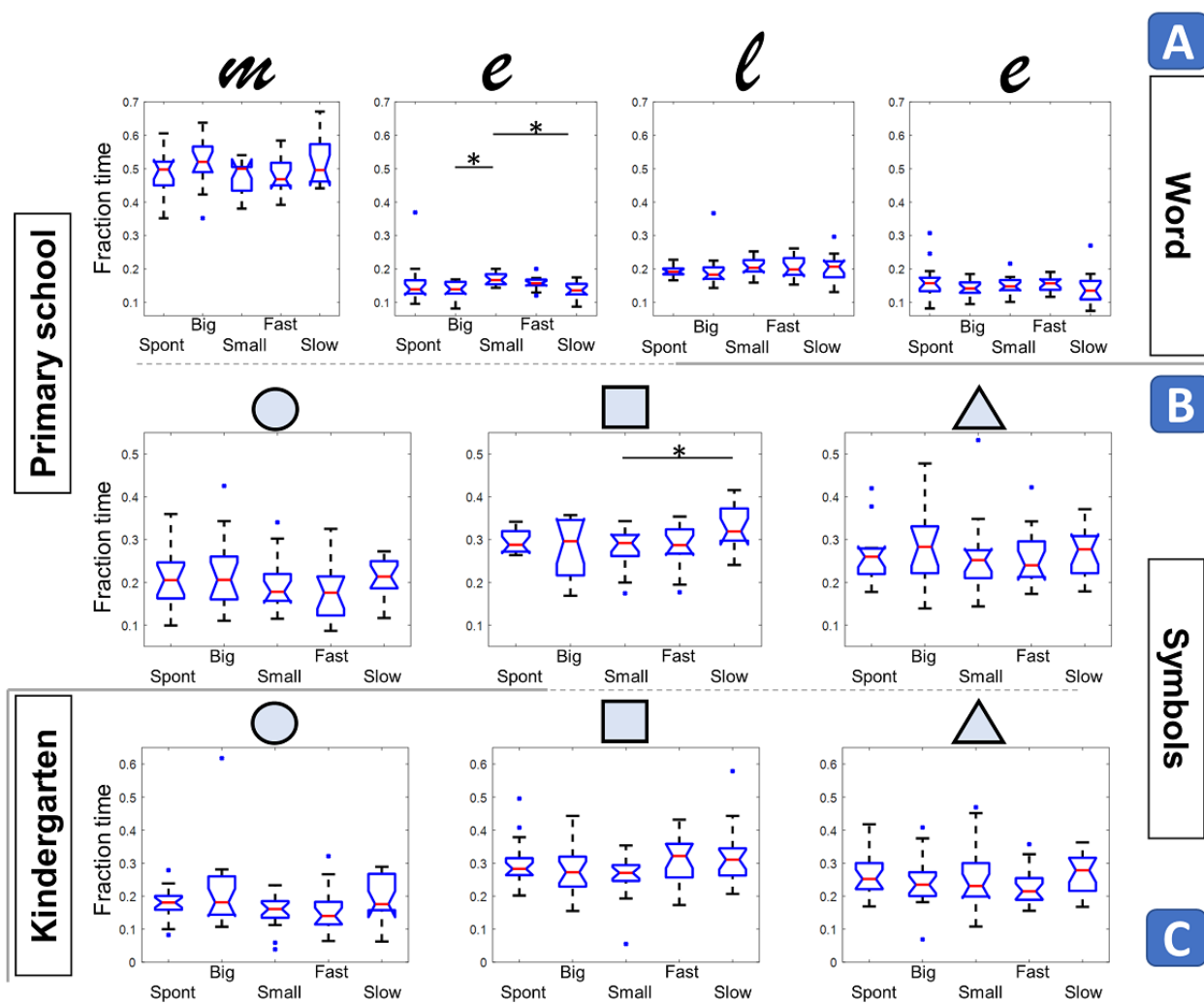
Figure 7. Speed-accuracy tradeoff. A: global movement time (ie, the median movement time computed across all indexes of difficulty). B: root mean square error (RMSE) (ie, the goodness of fit of the steering law). C: index of performance (ie, the inverse of the regression line slope). Data are divided by symbol/word and age group. *P<.05, **P<.005, ***P<.001.
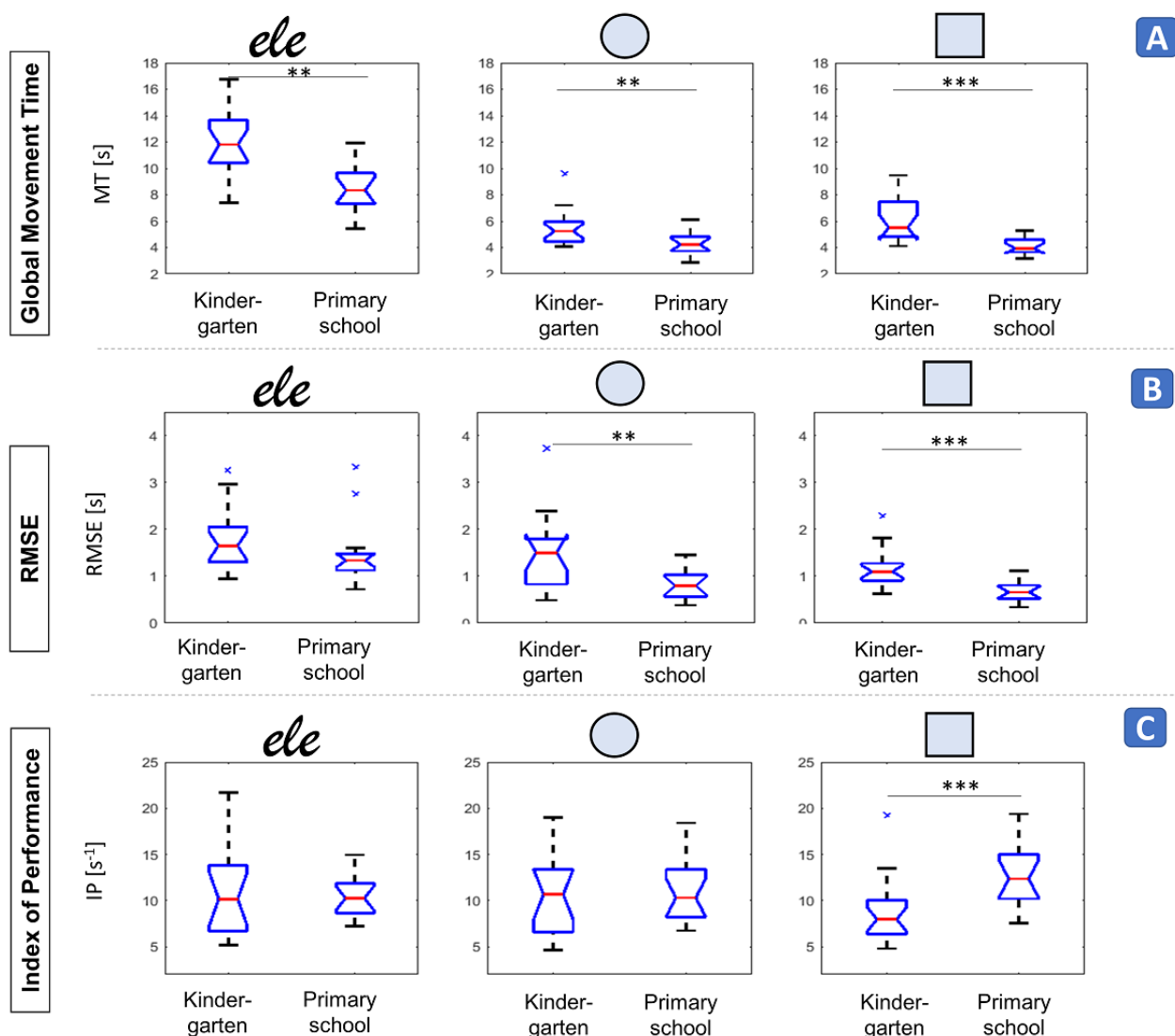
Copy Game

Concerning the copy game, instruction compliance was confirmed by all subjects in all the tests. Indeed, the trajectory in the *big* modality was always significantly longer than that traced in the *small* modality (P<.001, Multimedia Appendix 4, *exercise compliance*).

The speed execution of the word "mele" (Figure 5A; Multimedia Appendix 4, *isochrony*) reported a significant effect of the modality (P<.001) in primary school children. In particular, speed in the *big* modality was greater than in the *small* (P<.001) and *slow* modalities (P<.001), and that in the *fast* modality was greater than in the *small* (P<.001) and *slow* (P<.001) modalities.

A significant effect of the modality was also found for symbol speed execution in primary school children (P<.001) (Figure 5B). The speed in the *big* modality was greater than in the *small* (P<.001 in all cases) and *slow* (P=.002 for triangles, P<.001 in all other cases) modalities, and that in the *fast* modality was

greater than in the *small* (P<.001 in all cases) and *slow* (P<.001 in all cases) modalities.

The copy game (Figure 5C) showed a significant effect of the modality on speed for all symbols (P<.001) and for kindergartners. In particular, speed in the *big* modality was greater than in the *small* (P<.001 in all cases) and *slow* (P<.001 in all cases) modalities, and that in the *fast* modality was greater than in the *small* (P<.001 in all cases) and *slow* (P<.001 in all cases) modalities.

When investigating possible mean speed differences in the *spontaneous* modality between the two age groups, the Mann-Whitney U test did not reveal a significant effect of age for any of the symbols (circles: P=.55; triangles: P=.10; squares: P=.44; sequences: P=.50; Multimedia Appendix 4, *developmental trend*).

When analyzing the fraction time of letters in primary school children (Figure 6A; Multimedia Appendix 4, *homothety*), no significant between-condition differences were highlighted for M (P=.16), L (P=.20), and the last E (P=.21), thus supporting

the principle of homothety. A significant effect of the modality emerged only for the first *E* letter (P=.004). Post-hoc tests revealed that the fraction time for the first *E* letter in the *small* modality was significantly higher than in the *big* (P=.006) and *slow* (P=.009) modalities.

As for the fraction time in the sequence of symbols (Figure 6B), one primary school child was removed from the *big* modality dataset owing to erroneous execution. No fraction time differences emerged, with the only exception of squares between the *small* and *slow* modalities (P=.03).

As for the sequence of symbols in the kindergartners (Figure 6C), the modality had no effect on fraction time for squares (P=.06) and triangles (P=.30). We found a significant effect of the modality for circles (P=.02), which was not supported by the post-hoc assessment.

Tunnel Game

Among all trials, two kindergartners were not able to comply with the instructions during word steering, and thus, their data were removed.

Global MT (Figure 7A; Multimedia Appendix 5, *developmental trend*) was always significantly higher in kindergartners than in primary school children (word: P=.001; circles: P=.003; squares: $P<.001$).

In the primary school group, the linear fitting between ID and MT was significant for all children, except one in the square task. In the kindergarten group, all children obtained significant fitting for the square, whereas 15/19 circles and 15/17 word fittings were significant ($P<.05$, $R^2>0.28$, 12 degrees of freedom; Multimedia Appendix 5, *exercise compliance*).

The RMSE of symbol fitting (Figure 7B; Multimedia Appendix 5, *developmental trend*) was significantly lower (better fitting) for primary school children than for kindergartners (circle: P=.02; square: $P<.001$). The between-age difference in the RMSE of word fitting did not reach significance (P=.06), but the same trend was visible. RMSE differences between the word and symbols emerged in the primary school group ($P<.001$; Multimedia Appendix 5, *motor strategy differences*), but not in the kindergarten group (P=.08). In the primary school group, words fitted significantly worse than circles (P=.001) and squares ($P<.001$), but we did not find any significant difference between circles and squares (P=.65).

The IP (Figure 7C) showed a significant effect of age for squares ($P<.001$), but not circles (P=.65) or words (P=.80). No differences between the word and symbols emerged (primary school: P=.08; kindergarten: P=.13).

Discussion

A tablet-based app to investigate handwriting difficulties through serious games was developed and validated. Standard consumer technology, which does not require supervision and is suitable for home or school use, was leveraged to increase the usability and accessibility of the devised tool. The app was designed to quantitatively characterize the three principles underlying handwriting of isochrony, homothety, and speed-accuracy

tradeoff in not only word writing but also symbol drawing, with the final aim of anticipating the screening procedure for handwriting alterations in the preliteracy stage.

To verify that isochrony, homothety, and speed-accuracy tradeoff hold on a tablet surface during word writing, the app was tested on 15 typically developing third graders. Results showed that the tablet surface did not affect the presence of these laws for cursive word writing.

Specifically, isochrony was verified, as writing bigger always increases speed and writing smaller has the opposite effect. This confirms the results reported by Pagliarini et al [13], where 39 typically developing primary school children showed isochrony when writing on a piece of paper affixed on a professional digitizer. Our results extend the application domain of their findings, showing that a change in the writing support from a professional digitizer to a commercial tablet does not affect the presence of the isochrony law.

Additionally, the homothety principle holds on the tablet surface, as it was respected in 38 out of 40 cases during cursive writing. Neither writing size nor speed showed a relevant effect on the percentage of time dedicated to each letter with respect to the total word duration, with the only exception of one letter in the *small* modality, which significantly differed from both the *big* and *slow* modalities. This small exception may be ascribed to the nature of the cursive script. As letters are smoothly joined, different segmentation techniques might bring different outcomes. Indeed, our results on the tablet are in line with and further enhance previous work [13], which reported a higher number of exceptions of the homothety principle for the cursive script on a digitizer.

Finally, the speed-accuracy tradeoff principle was confirmed while writing on the tablet surface, since the steering law on words reached significance for 14 out of 15 primary school children. Importantly, these results support our proposal to leverage the steering law to study the speed-accuracy tradeoff during handwriting, which was never performed previously.

The above results support the use of tablet technology to quantitatively assess handwriting production. This is an important achievement, as previous research revealed that changing the writing surface is challenging for writers. Although literature reports an increase in writing speed resulting from the low friction coefficient of the tablet surface [24], this work proves that possible differences in specific parameters are not reflected in the violation of more high-level handwriting laws.

To detect possible handwriting alterations at an early stage, we investigated whether the principles underlying word writing hold also for symbol drawing in 15 typically developing third graders. Concerning isochrony, we found that it held for symbols. Indeed, isochrony-related speed modulation was observed for all symbols.

Results on homothety further support the possibility to replace a word with a sequence of symbols. Indeed, homothety regarding symbols was confirmed in 29 out of 30 cases. Our results showed that the principle of homothety is even stronger for a sequence of symbols than for a cursive word. An analogy can be found with the work of Pagliarini et al [13], where homothety

was verified for the capital script, but showed some exceptions for the cursive script. We argue that a sequence of capital script letters and a sequence of symbols are similar from both a planning and an execution point of view, thus making the two exercises interchangeable.

As for the speed-accuracy tradeoff for symbols, the linear regression of MT on ID was found to be relevant in 29 out of 30 trials.

Thus, we can conclude that symbols are valid candidates to substitute words and anticipate the study of possible violations of the principles, typical of dysgraphia.

Finally, we tested our app among 19 kindergartners to investigate if and how handwriting maturation affects the skills and laws underlying handwriting. Speed modulation and fraction time preservation were confirmed among kindergarteners during symbol drawing in 100% of cases, supporting that adherence to the isochrony and homothety principles when writing on a tablet surface does not develop as a consequence of mastering handwriting [17-19]. Previous work [19] reported some exceptions to the homothety principle during block letter writing, especially in younger children. In light of our results, we claim that a sequence of symbols is more suitable than a block letter word when investigating homothety at a preliteracy age. Indeed, symbol drawing has the advantages of being acquired years before handwriting and being language independent [32].

Among kindergartners, the speed-accuracy tradeoff was studied while steering both symbols and a cursive word, since we hypothesized that path guidance was a valid help in task execution even before writing acquisition. Our findings showed that the linear regression of MT on ID was significant in the majority of cases ($P<.05$, $R^2>0.28$, 12 degrees of freedom, 15/17 for words and 34/38 for symbols). Our results support that this principle is innate both in path steering and in handwriting, and can be potentially used as a screening tool at a preliteracy age.

The steering law parameters revealed a developing trend with age. In particular, kindergartners needed more time to complete the speed-accuracy exercises than older children. This result seems to contradict the group invariance of the average speed obtained when copying symbols (copy game in the *spontaneous* modality). We can conjecture that sticking to complex tasks is more demanding for kindergarteners, but they do not have any specific difficulty in free execution. This consideration is further supported when analyzing the IP during square tracing. Kindergartners achieved significantly lower results than primary school children ($P<.001$), showing that they need a bigger adaptation effort (ie, longer MT) to react to different difficulty

constraints. The square symbol seems to be the most adequate to disclose such between-group difference, probably since it is the most similar to the steering law original formulation [15]. The same age-related trend was found in previous work on the speed-accuracy tradeoff in the discrete domain [30,33], supporting the use of the speed-accuracy tradeoff principle as an index of maturity of motion control. The same law was also proven to successfully discriminate between healthy and pathological conditions (eg, dystonia [34], and Huntington and Parkinson diseases [35]). In this work, we further proved that steering's IP is useful to differentiate between ages, and we speculate that it has the potential to differentiate between pathological and normal handwriting as well. Therefore, we plan to continue our experiments on a group of pupils with dysgraphia to further prove the opportunity to exploit our tool to support the identification of handwriting abnormalities.

Finally, a higher RMSE emerged for kindergartners during symbol drawing, suggesting that adherence to the steering law is a process that is refined with age. In primary school children only, the RMSE for word tracing was significantly higher than that for symbol tracing ($P=.001$ for word-circle comparison, $P<.001$ for word-square comparison). This result can be explained by the different kinds of movements required to produce a letter. Indeed, mastered handwriting can be considered open loop [36], and this aspect might alter the kinematic pattern modelled by the steering law. To support such a hypothesis, the lack of differences in kindergartners suggests that they approached words and symbols similarly, as they did not learn or automate handwriting yet. Further experiments can clarify this aspect.

In conclusion, we proposed a simple tool to quantify the degree of deviation from physiological handwriting, which could help with the objective detection of dysgraphic alterations. The tool exploits serious games to assess important rhythmic principles of handwriting, such as isochrony, homothety, and speed-accuracy tradeoff. It does not require any professional equipment or professional supervision, making the assessment simple and accessible in nonclinical environments. Importantly, we proved that it is possible to start evaluating characteristics underlying handwriting in preschoolers through the use of symbols, which makes the evaluation language independent. Eventually, the tool can provide interesting insights on graphomotor abilities correlated with handwriting and can be used to monitor their evolution over time. Thus, the proposed tool paves the way toward a new model of care. It includes the screening of preliteracy skills to identify potential difficulties before they arise and offers continuous, accessible, and low-cost monitoring of their evolution.

Conflicts of Interest

None declared.

XSL•FO

RenderX

Multimedia Appendix 1

Copy game: sequence.

[MP4 File (MP4 Video), 156574 KB - games_v8i4e20126_app1.mp4]

Multimedia Appendix 2

Tunnel game: word.

[MP4 File (MP4 Video), 33699 KB - games_v8i4e20126_app2.mp4]

Multimedia Appendix 3

Index of difficulty calculation.

[PDF File (Adobe PDF File), 113 KB - games_v8i4e20126_app3.pdf]

Multimedia Appendix 4

Statistical test results for the copy game.

[PDF File (Adobe PDF File), 179 KB - games_v8i4e20126_app4.pdf]

Multimedia Appendix 5

Statistical test results for the tunnel game.

[PDF File (Adobe PDF File), 147 KB - games_v8i4e20126_app5.pdf]

References

1. Chung PJ, Patel DR, Nizami I. Disorder of written expression and dysgraphia: definition, diagnosis, and management. Transl Pediatr 2020 Feb;9(Suppl 1):S46-S54 [FREE Full text] [doi: 10.21037/tp.2019.11.01] [Medline: 32206583]

2. Prunty MM, Barnett AL, Wilmut K, Plumb MS. Handwriting speed in children with Developmental Coordination Disorder: are they really slower? Res Dev Disabil 2013 Sep;34(9):2927-2936 [FREE Full text] [doi: 10.1016/j.ridd.2013.06.005] [Medline: 23816628]

3. Asselborn T, Chapatte M, Dillenbourg P. Extending the Spectrum of Dysgraphia: A Data Driven Strategy to Estimate Handwriting Quality. Sci Rep 2020 Feb 21;10(1):3140 [FREE Full text] [doi: 10.1038/s41598-020-60011-8] [Medline: 32081940]

4. Feder KP, Majnemer A. Handwriting development, competency, and intervention. Dev Med Child Neurol 2007 Apr;49(4):312-317 [FREE Full text] [doi: 10.1111/j.1469-8749.2007.00312.x] [Medline: 17376144]

5. Battle DE. Diagnostic and Statistical Manual of Mental Disorders (DSM). Codas 2013;25(2):191-192. [doi: 10.1590/s2317-17822013000200017] [Medline: 24413388]

6. Al-Yagon M, Cavendish W, Cornoldi C, Fawcett AJ, Grünke M, Hung L, et al. The proposed changes for DSM-5 for SLD and ADHD: international perspectives—Australia, Germany, Greece, India, Israel, Italy, Spain, Taiwan, United Kingdom, and United States. J Learn Disabil 2013;46(1):58-72 [FREE Full text] [doi: 10.1177/0022219412464353] [Medline: 23197670]

7. Overvelde A, Hulstijn W. Handwriting development in grade 2 and grade 3 primary school children with normal, at risk, or dysgraphic characteristics. Res Dev Disabil 2011;32(2):540-548 [FREE Full text] [doi: 10.1016/j.ridd.2010.12.027] [Medline: 21269805]

8. Karlsdottir R, Stefansson T. Problems in developing functional handwriting. Percept Mot Skills 2002 Apr;94(2):623-662 [FREE Full text] [doi: 10.2466/pms.2002.94.2.623] [Medline: 12027360]

9. Mather DS. Dyslexia and dysgraphia: more than written language difficulties in common. J Learn Disabil 2003;36(4):307-317 [FREE Full text] [doi: 10.1177/00222194030360040301] [Medline: 15490904]

10. Ben-Pazi H, Kukke S, Sanger TD. Poor penmanship in children correlates with abnormal rhythmic tapping: a broad functional temporal impairment. J Child Neurol 2007 May;22(5):543-549 [FREE Full text] [doi: 10.1177/0883073807302610] [Medline: 17690059]

11. Viviani P, Terzuolo C. Trajectory determines movement dynamics. Neuroscience 1982 Feb;7(2):431-437 [FREE Full text] [doi: 10.1016/0306-4522(82)90277-9] [Medline: 7078732]

12. Lashley KS. The problem of serial order in behavior. In: Jeffress LA, editor. Cerebral Mechanisms in Behavior. New York: John Wiley & Sons; 1951:112-136.

13. Pagliarini E, Guasti MT, Toneatto C, Granocchio E, Riva F, Sarti D, et al. Dyslexic children fail to comply with the rhythmic constraints of handwriting. Hum Mov Sci 2015 Aug;42:161-182 [FREE Full text] [doi: 10.1016/j.humov.2015.04.012] [Medline: 26037277]

14. Smits-Engelsman BC, Wilson PH, Westenberg Y, Duysens J. Fine motor deficiencies in children with developmental coordination disorder and learning disabilities: an underlying open-loop control deficit. Hum Mov Sci 2003 Nov;22(4-5):495-513 [FREE Full text] [doi: 10.1016/j.humov.2003.09.006] [Medline: 14624830]

XSL•FO

RenderX

15.   Accot J, Zhai S. Beyond Fitts' law: models for trajectory-based HCI tasks. In: CHI EA '97: CHI '97 Extended Abstracts on Human Factors in Computing Systems. 1997 Presented at: ACM SIGCHI Conference on Human Factors in Computing Systems; 1997; Atlanta p. 250. [doi: 10.1145/1120212.1120376]

16.   Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology: General 1992 Sep;121(3):262-269 [FREE Full text] [doi: 10.1037//0096-3445.121.3.262] [Medline: 1402698]

17.   Vinter A, Mounoud P. Isochrony and accuracy of drawing movements in children: Effects of age and context. In: Wann J, Wing AM, Sõvik N, editors. Development of Graphic Skills: Research Perspectives and Educational Implications. Cambridge, Massachusetts: Academic Press; 1991:113-134.

18.   Viviani P, Schneider R. A developmental study of the relationship between geometry and kinematics in drawing movements. Journal of Experimental Psychology: Human Perception and Performance 1991 Feb;17(1):198-218 [FREE Full text] [doi: 10.1037//0096-1523.17.1.198] [Medline: 1826312]

19.   Pagliarini E, Scocchia L, Vernice M, Zoppello M, Balottin U, Bouamama S, et al. Children's first handwriting productions show a rhythmic structure. Sci Rep 2017 Jul 17;7(1):5516 [FREE Full text] [doi: 10.1038/s41598-017-05105-6] [Medline: 28717141]

20.   Lunardini F, Luperto M, Romeo M, Basilico N, Daniele K, Azzolino D, et al. Supervised Digital Neuropsychological Tests for Cognitive Decline in Older Adults: Usability and Clinical Validity Study. JMIR Mhealth Uhealth 2020 Sep 21;8(9):e17963 [FREE Full text] [doi: 10.2196/17963] [Medline: 32955442]

21.   Howe T, Chen H, Lee CC, Chen Y, Wang T. The Computerized Perceptual Motor Skills Assessment: A new visual perceptual motor skills evaluation tool for children in early elementary grades. Res Dev Disabil 2017 Oct;69:30-38 [FREE Full text] [doi: 10.1016/j.ridd.2017.07.010] [Medline: 28802160]

22.   Bang C, Nam Y, Ko EJ, Lee W, Kim B, Choi Y, et al. A Serious Game-Derived Index for Detecting Children With Heterogeneous Developmental Disabilities: Randomized Controlled Trial. JMIR Serious Games 2019 Oct 24;7(4):e14924 [FREE Full text] [doi: 10.2196/14924] [Medline: 31651408]

23.   Borghese NA, Palmiotto C, Essenziale J, Mainetti R, Granocchio E, Molteni B, et al. Assessment of Exergames as Treatment and Prevention of Dysgraphia. 2017 Presented at: Converging Clinical and Engineering Research on Neurorehabilitation II; October 2016; Segovia p. 431-436. [doi: 10.1007/978-3-319-46669-9_72]

24.   Gerth S, Klassert A, Dolk T, Fliesser M, Fischer MH, Nottbusch G, et al. Is Handwriting Performance Affected by the Writing Surface? Comparing Preschoolers', Second Graders', and Adults' Writing Performance on a Tablet vs. Paper. Front Psychol 2016;7:1308 [FREE Full text] [doi: 10.3389/fpsyg.2016.01308] [Medline: 27672372]

25.   Verschueren S, Buffel C, Vander Stichele G. Developing Theory-Driven, Evidence-Based Serious Games for Health: Framework Based on Research Community Insights. JMIR Serious Games 2019 May 02;7(2):e11565 [FREE Full text] [doi: 10.2196/11565] [Medline: 31045496]

26.   Pirovano M, Surer E, Mainetti R, Lanzi P, Alberto Borghese N. Exergaming and rehabilitation: A methodology for the design of effective and safe therapeutic exergames. Entertainment Computing 2016 May;14:55-65 [FREE Full text] [doi: 10.1016/j.entcom.2015.10.002]

27.   Van Galen GP, Van Doorn RR, Schomaker LR. Effects of motor programming on the power spectral density function of finger and wrist movements. Journal of Experimental Psychology: Human Perception and Performance 1990 Nov;16(4):755-765 [FREE Full text] [doi: 10.1037//0096-1523.16.4.755] [Medline: 2148590]

28.   Tressoldi PE, Cornoldi C. Batteria per la valutazione della scrittura e della competenza ortografica nella scuola dell'obbligo: manuale. Florence, Italy: Organizzazioni speciali; 1991.

29.   Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. Pediatrics 1992 Jan;89(1):91-97. [Medline: 1370185]

30.   Hourcade JP, Bederson BB, Druin A, Guimbretière F. Differences in pointing task performance between preschool children and adults using mice. ACM Trans. Comput.-Hum. Interact 2004 Dec;11(4):357-386 [FREE Full text] [doi: 10.1145/1035575.1035577]

31.   MacKenzie IS, Isokoski P. Fitts' throughput and the speed-accuracy tradeoff. In: Proceedings of the ACM Conference on Human Factors in Computing Systems – CHI 2008. 2008 Presented at: ACM Conference on Human Factors in Computing Systems; April 2008; Florence, Italy p. 1633-1636 URL: https://dl.acm.org/doi/abs/10.1145/1357054. 1357308?casa_token=k4QujLhGs60AAAAA:czEqBJb3RyWt3OoytAjA0AiS_lZLPu5aJCZ6KxtyWeHkdhM9LhF4t8wkZNvZv0txcpjOA3lXz2yo6A

32.   Evertz M, Primus B. Rhythmic constraints in handwriting. A review of Pagliarini et al. (2017). 2017 Dec 31;20(2):252-258. [doi: 10.1075/wll.00006.eve]

33.   Smits-Engelsman BC, Sugden D, Duysens J. Developmental trends in speed accuracy trade-off in 6-10-year-old children performing rapid reciprocal and discrete aiming movements. Hum Mov Sci 2006 Feb;25(1):37-49 [FREE Full text] [doi: 10.1016/j.humov.2005.12.002] [Medline: 16442174]

34.   Lunardini F, Bertucco M, Casellato C, Bhanpuri N, Pedrocchi A, Sanger TD. Speed-Accuracy Trade-Off in a Trajectory-Constrained Self-Feeding Task: A Quantitative Index of Unsuppressed Motor Noise in Children With Dystonia. J Child Neurol 2015 Oct;30(12):1676-1685 [FREE Full text] [doi: 10.1177/0883073815578526] [Medline: 25895910]

35. Halsband U, Homberg V, Lange HJ. Slowing of different types of voluntary movement in extrapyramidal disease: Fitts' law and idiographic writing. In: Barardelli A, editor. Motor Disturbances II. New York: Plenum Press; 1990:181-190.

36. Meulenbroek RG, Van Galen GP. The acquisition of skilled handwriting: Discontinuous trends in kinematic variables. In: Colley AM, Beech JR, editors. Advances in psychology: Vol. 55. Cognition and action in skilled behaviour. Holland: Elsevier; 1988:273-281.

Abbreviations

CI: confidence interval
ID: index of difficulty
IP: index of performance
IQR: interquartile range
MT: movement time
RMSE: root mean square error

XSL•FO
RenderX

Original Paper

Evaluation of a Serious Video Game to Facilitate Conversations About Human Papillomavirus Vaccination for Preteens: Pilot Randomized Controlled Trial

Joan R Cates[1*], BA, MPH, PhD; Bernard F Fuemmeler[2*], MPH, PhD; Laurie L Stockton[1*], MPH; Sandra J Diehl[1*], MPH; Jamie L Crandell[1*], PhD; Tamera Coyne-Beasley[3], MD, MPH

[1]UNC Hussman School of Journalism and Media, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]Massey Cancer Center, Virginia Commonwealth University, Richmond, VA, United States

[3]Department of Pediatrics, Children's of Alabama, University of Alabama, Birmingham, AL, United States

[*]these authors contributed equally

Corresponding Author:
Bernard F Fuemmeler, MPH, PhD
Massey Cancer Center
Virginia Commonwealth University
830 E. Main St., Rm 4-128
P.O. Box 980149
Richmond, VA, 23219
United States
Phone: 1 804 828 8892
Email: Bernard.Fuemmeler@vcuhealth.org

Abstract

Background: In the United States, the most common sexually transmitted infection, human papillomavirus (HPV), causes genital warts and is associated with an estimated 33,700 newly diagnosed cancer cases annually. HPV vaccination, especially for preteens aged 11 to 12 years, is effective in preventing the acquisition of HPV and HPV-associated cancers. However, as of 2018, completion of the 2- or 3-dose HPV vaccination series increased only from 48.6% to 51.1% in teens aged 13 to 17 years, and this increase was observed only in boys. By comparison, 88.7% of teens had more than one dose of the recommended vaccine against tetanus, diphtheria, and acellular pertussis (Tdap), and 85.1% of teens had more than one dose of meningococcal vaccine. Immunizations for Tdap, meningococcal disease, and HPV can occur at the same clinical visit but often do not.

Objective: Vaccination against HPV is recommended for routine use in those aged 11 to 12 years in the United States, yet it is underutilized. We aimed to develop an educational video game to engage preteens in the decision to vaccinate.

Methods: *Land of Secret Gardens* is a metaphor for protecting seedlings (body) with a potion (vaccine). We screened 131 dyads of parents and preteens from 18 primary practices in North Carolina who had not initiated HPV vaccination. We measured vaccination intentions, story immersion, and game play and documented HPV vaccination rates. A total of 55 dyads were enrolled, and we randomly assigned 28 (21 completed) to play the game and 27 (26 completed) to the comparison group.

Results: In total, 18 preteens reported playing the game. The vaccination self-efficacy score was higher in the comparison group than the intervention group (1.65 vs 1.45; *P*=.05). The overall mean decisional balance score trended toward greater support of vaccination, although differences between the groups were not significant.. Vaccine initiation and completion rates were higher in the intervention group (22% vs 15%; *P*=.31) than in the comparison group (9% vs 2%; *P*=.10), although the difference was not significant.

Conclusions: Video games help preteens in the decision to pursue HPV vaccination. A serious video game on HPV vaccination is acceptable to parents and preteens and can be played as intended. Gamification is effective in increasing preteen interest in HPV vaccination, as game features support decision making for HPV vaccination.

Trial Registration: ClinicalTrials.gov NCT04627298; https://www.clinicaltrials.gov/ct2/show/NCT04627298

XSL•FO
RenderX

Introduction

Background

More than a decade has passed since a vaccine to prevent human papillomavirus (HPV) infection was recommended for routine use in children aged 11 to 12 years in the United States [1]. The most common sexually transmitted infection in the United States, HPV, causes genital warts and is associated with an estimated 33,700 newly diagnosed cancer cases [2]. However, as of 2018, completion of the 2- or 3-dose HPV vaccination series increased only from 48.6% to 51.1% in teens aged 13 to 17 years, and this increase was observed only in boys [1]. By comparison, 88.7% of teens had more than 1 dose of the recommended vaccine against tetanus, diphtheria, and acellular pertussis (Tdap), and 85.1% of teens had more than 1 dose of meningococcal vaccine [2]. Immunizations for Tdap, meningococcal disease, and HPV can occur at the same clinical visit but often do not [3]. Testing and evaluating practice-based implementation strategies are needed to improve the uptake of effective interventions to increase HPV vaccination initiation and completion.

Full HPV vaccination coverage has been challenging due, in part, to providers not making strong recommendations [1]. There also remain parental concerns about the vaccine. For instance, some parents perceived the risk of HPV infection to be negligible, expressed concern about side effects, and believed the vaccine might encourage promiscuous behavior or that it may be too costly [4,5]. Health care professionals have reported that these parental attitudes and concerns are barriers to vaccination [4]. Thus, understanding and addressing these barriers would be critical to target within the context of interventions designed to increase uptake.

Although many interventions promoting HPV vaccination have focused on either the parent or provider separately with moderate success [6,7], it is increasingly being recognized that a multilevel approach may further broaden dissemination efforts [8,9]. Communication strategies have focused on giving accurate information about HPV vaccination and on training providers to give clear messages about the safety and efficacy of HPV vaccines [10].

In response to this challenge of suboptimal HPV vaccination, our interdisciplinary team (communication, public health, medicine, clinical psychology, biostatistics, health economics, and nursing) has been leading efforts to design and implement multilevel communication strategies that target parents, health care providers, and preteens [11-13]. For instance, our approach has been to develop methods to address parent resistance and misunderstandings about why the vaccine is needed early in development before their children are sexually active [14,15]. In addition, we have worked to develop strategies to address provider's perceived barriers about discussing the vaccine with their patients (eg, helping them to develop *talking points* that can be used within a short patient visit) [16]. Finally, recognizing that serious video games may have the potential to

educate preteens about sexually transmitted diseases or the utility of vaccination [17,18] and can be effective in promoting health behaviors in children and adolescents [19-23], we developed a serious video game to promote HPV vaccination among preteens [24]. Specifically, with input from preteens aged 11 to 12 years and their parents, we developed *Land of Secret Gardens*, a serious video game designed to teach about vaccines through an immersive story and to motivate a decision to seek HPV vaccination [24].

Putting all of these pieces together, we initiated the Protect Them study, which was undertaken in 36 primary care practices with 97 providers (MD, DO, NP, and RN) in North Carolina. This was a multiple baseline study and included 3 waves of activity and adjustment in 2015, 2016, and 2017. The intervention was designed to promote communication among providers, parents, and preteens to increase HPV vaccination for preteens aged 11 to 12 years when the vaccine is most effective [25]. Communication tools for the providers included brochures, posters, web-based information for parents, and interactive web-based training for providers. In addition, as part of the intervention, we provided select patients access to the Land of the Secret Garden, an age-appropriate, entertaining web-based video game designed to educate preteens about HPV infection and HPV vaccination and to promote conversations with parents and providers and the decision to vaccinate.

The video game incorporated gamification elements (eg, points, badges for completion of tasks, and a leaderboard) [26] to increase interest among players while also aiming to increase preteen knowledge and vaccine self-efficacy (ie, confidence in getting the vaccine despite barriers). Self-determination theory [27] was also used to inform the game design, as this has previously been used to evaluate the motivational pull of video games [23]. In addition, the game included an immersive story to enhance motivation to play the game [23,28] and engender deeper information processing [21]. Our hypothesis was that raising awareness about HPV vaccination eases conversations about the vaccination. We created a story about a secret garden as a metaphor for a preteen's body and keeping it healthy. The goal was to plant a lush secret garden and protect the seedlings by treating them with a potion when they sprout to keep them healthy as they mature. Points to buy seeds and create the potion were earned by playing minigames. The minigames included several versions of finding secret objects in a garden shed and another that involved shooting down spikey balls (ie, the HPV) before they land on budding plants. Throughout the play, players were exposed to messaging about HPV and the benefits of the vaccine [24].

Objectives

Herein, we report on the evaluation of the *Land of Secret Gardens* game. The aims of this pilot study are to evaluate (1) preliminary data to determine whether children who received the *Land of Secret Gardens* game had better self-reported outcomes related to HPV knowledge or vaccination self-efficacy compared with those in a control group who did not receive the

XSL·FO

RenderX

game and (2) outcomes related to the game play experience (in-game autonomy and competence, presence in the game, intuitive controls, interest or enjoyment, and characteristics of playing the game) among those who received the game. We also conducted focus groups among those who received the game to further assess the acceptability of the game and whether the preteens understood the meaning of the game. Finally, we compared HPV vaccination initiation and completion among those who received the game compared with the control group who did not receive the game.

Methods

Participants and Procedures

Participants of this video game evaluation study were part of the lager Protect Them study mentioned earlier in the Introduction section. Parent and preteen dyads were recruited by providers at 36 different clinical sites that included family medicine practices, pediatric practices, and health departments. To enroll in the study, clinical sites in the Protect Them study agreed to identify up to 10 parents (wave 1) of preteens aged 11 to 12 years and up to 20 parents (waves 2 and 3) of preteens. Preteens aged 11 to 12 years and who had not received any doses of HPV vaccine were eligible to participate in the study. Providers contacted potential dyads via letters or telephone after they identified eligible preteens from their electronic medical records. Interested dyads provided their contact information (ie, phone number and email address) to the research team and signed a Health Insurance Portability and Accountability Act (HIPAA) release of information form to allow the research team to determine their HPV vaccination status from their health care provider. A practice champion provided a copy of the HIPAA form and the parent contact information to project staff. The most common obstacle was a high proportion of clients with parents needing informed consent provided in Spanish, an accommodation not available for this study.

After the names of eligible parents and preteens were passed to the research team, research staff invited parents to participate in a telephone conversation about informed consent with their preteen via email and telephone contacts. The study protocol required up to 3 attempts to reach parents by both email and telephone contact. Staff used an institutional review board (IRB)–approved script to conduct an informed consent procedure with the parent and then with the preteen if the parent provided consent. The script included a description of the intervention and the process of random assignment to a study group, explained the risks and benefits of participation, and reviewed the study activities and incentives for participation. Both parents

and preteens were assured that their information was confidential and that participation was voluntary. In addition to preteens who had not received any HPV vaccine, eligible dyads confirmed that they had access to the internet and a mobile device or personal computer to complete the surveys and play the video game. They also provided a mailing address to receive gift certificates. Once they were enrolled, dyads were randomly assigned in a 1:1 ratio to either the intervention group or control group using a simple randomization schedule generated by the study team's statistician. Those who were assigned to the intervention arm received the game, whereas those assigned to the control group arm did not receive the game. The study procedures were approved by the university's nonbiomedical IRB.

A total of 36 practices across 3 regions in North Carolina enrolled in the Protect Them study (Figure 1)—12 practices were recruited from the eastern region during wave 1 (February 2015 to June 2015), 18 practices were recruited from the central region during wave 2 (February 2016 to July 2016), and 6 practices were recruited from the western region during wave 3 (February 2017 to October 2017).

All practices were asked to screen dyads as part of the intervention. Half of the practices (18/36, 50%) screened at least one dyad, and a total of 131 dyads were referred to the research team. Of these 131 dyads, 16 (12.2%) did not meet the eligibility criteria for the study. Among the 115 that were eligible, 16 (13.9%) refused to participate in the study, and 42 (36.5%) dyads did not respond to repeated email and telephone contact. Almost half of the eligible dyads (55/115, 47.8%) completed the informed consent process and baseline surveys and were enrolled in the study, and 47 dyads completed the postsurvey.

The final study sample available for this evaluation of the game consisted of 28 preteens in the intervention group and 27 in the comparison group (21 and 26, respectively, completed the study). Preteens in the intervention group were 57% (16/28) female and 71% (20/28) white, and in the comparison group, 52% (14/27) were female and 82% (22/27) were white. The comparison group included 5 Blacks or African Americans, and the intervention group included 4 Blacks or African Americans. Two participants in each group were identified as Hispanic. The intervention group contained 9 participants aged 11 years and 19 participants aged 12 years, and the comparison group contained 14 participants aged 11 years, 11 participants aged 12 years, and 2 participants aged 13 years. There were no statistically significant differences between the groups with respect to gender, age, race, and Hispanic ethnicity.

Figure 1. Sample results from recruitment of parent or preteen dyads through practices (n=36).
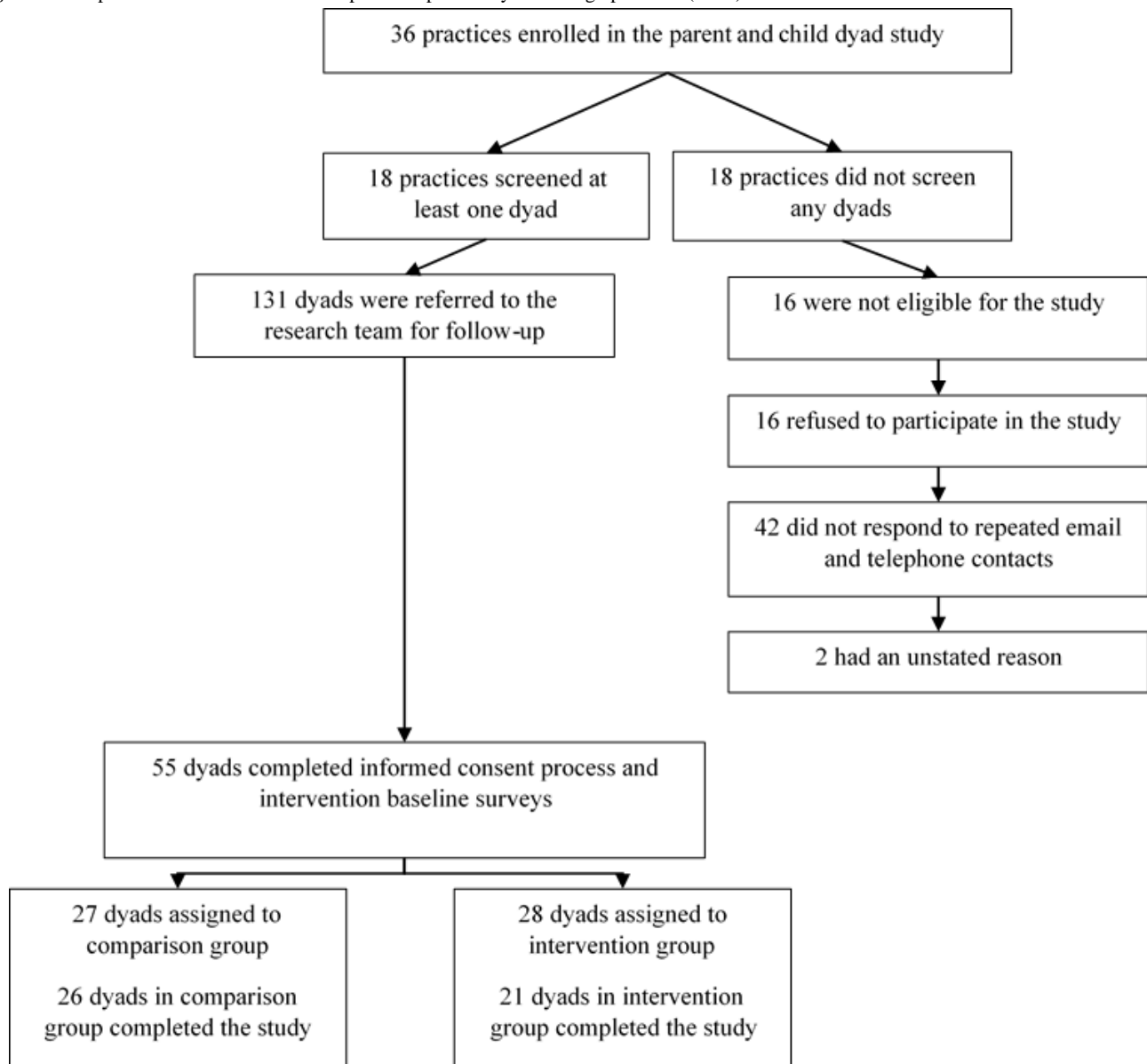
Survey Procedures

We asked all parents and preteens to complete the baseline and postintervention surveys. The surveys were designed in Qualtrics [29], and survey links were sent to the parent's email address. Dyads received the baseline surveys before potential exposure to the video game and parent portal. Up to 5 email reminders and 3 telephone calls were made to encourage survey completion. Follow-up reminders were offered to encourage game play and completion of the task. Each participant received a US $25 gift certificate to Walmart if they completed a survey.

Study groups were asked identical questions about their knowledge and attitudes about HPV and their intentions to vaccinate against HPV at baseline. Postintervention surveys were sent 4 months after participants completed their baseline surveys. All participants were asked if the preteen received any dose of the HPV vaccine. Participants in the intervention group were asked additional questions about their experience with the intervention and the Protect Them resources.

Game Play Procedures

Preteens in the intervention group were asked to play the *Land of Secret Gardens* and complete 3 tasks in the video game. The tasks occurred in a sequence that required the player to return to the game multiple times. A badge appeared on the leaderboard on completion of each task, which allowed preteens to track their progress. Preteens were exposed to messages about HPV and the HPV vaccination throughout the game, and continued use of the game would result in greater message exposure. Project staff monitored the game play progress for each participant with Navicat [30] and sent reminders to parents to encourage game play. Instructions for parents to guide game play were posted on the parent portal, along with a video that described the background and premise of the game. A help form was also available on the parent portal to request technical support for the game.

Measures

We assigned preteens in the intervention group to play the video game and asked their parents to review web-based materials

and then collected baseline and postintervention data from both study groups via Qualtrics surveys. We measured knowledge, vaccination self-efficacy, and decisional balance in both groups. From the intervention group only, we collected data on Physical/Emotional/Narrative Presence Scale (PENS) [31] to gauge preteens' immersion in the game. Finally, from the practice champion, we collected the HPV vaccination status of each preteen participant approximately 9 months postintervention.

The description of each measure is as follows:

- *Knowledge scale*: asked in both the intervention and comparison groups. The 5 items asked whether HPV vaccination (1) can prevent genital warts, (2) can prevent cervical cancer, (3) can prevent anal cancer, (4) can prevent throat cancer, and (5) is recommended for 11- and 12-year-old boys and girls.
- *Vaccination Self-efficacy and decisional balance scales* [32]: used in the intervention group and the comparison group to compare self-efficacy and intentions to vaccinate and decisional balance. We used 8 items for vaccination self-efficacy, 4 items for positive decision to vaccinate, and 5 items for negative decision to vaccinate. All questions were rated on a 3-point Likert scale (1=not at all, 2=somewhat, and 3=a lot).
- *PENS* [31]: to learn more about the intervention experience, we collected measures specific to preteens in the intervention group. We used measurements from PENS to gauge the extent of the preteen's immersion in the story.
- *Game play*: we asked participants who played the game how they played the game, for example, earning at least one shield, playing 3 or more times, playing more than 10 min per session, playing with a parent or sibling, playing the shield game and the hidden objects game, creating a potion, and correctly identifying the game metaphor. In addition, we recruited 3 postintervention focus groups with preteens (2 or 3 in each group) who played the game and

asked about their experience. The focus groups were conducted and recorded via telephone calls. Parents gave consent before the preteens joined the conversation. The study moderator asked whether the preteens enjoyed the game and which parts they enjoyed or did not enjoy.

- *HPV immunization records*: obtained for all preteen participants from practice champions approximately 9 months following each intervention. This period allowed enough interval for the preteens to complete their HPV vaccination series.

Statistical Analysis

We compared postintervention knowledge, vaccination self-efficacy, and decisional balance in the intervention and comparison groups using two-sample t tests, with significance level of α=.05. Immunization records were obtained for all preteen participants approximately 9 months postintervention for each cohort total. During the time of the study, the HPV vaccine was offered at both a 2-dose schedule with up to 6 to 12 months after the first shot and a 3-dose schedule with up to 6 months after the first shot. Practice champions were asked whether a preteen was on a 2- or 3-dose HPV vaccination schedule and to confirm whether they initiated and/or completed the vaccine series. The intervention and comparison groups were compared regarding the initiation and completion of the vaccine series using the Mantel-Haenszel chi-square test stratified by intervention wave.

Results

Knowledge and Vaccination Self-Efficacy and Decisional Balance

Postintervention, the mean knowledge score (5 items, range 1-3) was higher in the intervention group than in the comparison group (2.56, SD 0.34 vs 2.28, SD 0.41 respectively; *P*=.03). The results of the self-efficacy and decisional balance results are given in Table 1 and summarized as follows.

Table 1. Mean comparisons between items from the self-efficacy and decisional balance scales at post-intervention surveys.

Items from self-efficacy and decisional balance scales	Intervention (n=21), mean (SD)	Comparison (n=26), mean (SD)	P value
Vaccination self-efficacy. How confident am I about getting the vaccine…			
When I think about the possible side effects of the vaccine?	1.77 (0.69)	1.85 (0.61)	.70
When I think that the shot will be painful?	1.91 (0.75)	1.92 (0.74)	.95
When my parents are getting me vaccinated?	1.29 (0.56)	1.46 (0.65)	.33
When I think I will faint or get dizzy when getting the shot?	1.55 (0.74)	1.46 (0.71)	.69
When it is too expensive?	1.18 (0.5)	1.54 (0.65)	.04
When it is too inconvenient?	1.23 (0.53)	1.73 (0.60)	.01
When the doctor does not strongly recommend it?	1.41 (0.67)	1.62 (0.64)	.28
When my friends will know I got the shot?	1.23 (0.53)	1.62 (0.75)	.04
All vaccination self-efficacy items	1.45 (0.35)	1.65 (0.35)	.05
Importance. How important is this item in deciding to get HPV[a] vaccination?			
Pros			
Protecting myself from HPV would make me feel good.	2.55 (0.51)	2.39 (0.50)	.31
I would be protected from certain cancers and genital warts.	2.57 (0.51)	2.55 (0.51)	.87
I would be protecting myself from getting a sexually transmitted infection	2.36 (0.58)	2.31 (0.74)	.77
I would be less likely to spread HPV	2.19 (0.81)	1.92 (0.8)	.26
All pros items	2.44 (0.39)	2.31 (0.42)	.29
Cons			
Receiving the series would take too much time	1.77 (0.69)	1.73 (0.72)	.84
It would be too embarrassing to talk to my parents	1.41 (0.59)	1.31 (0.55)	.54
It would be too embarrassing to talk to my doctor about getting vaccinated	1.55 (0.6)	1.5 (0.65)	.80
My parents would not approve of me receiving the vaccine	1.1 (0.31)	1.62 (0.8)	.01
My parents would think I was having sex if I got vaccinated	1.0 (0.01)	1.19 (0.57)	.13
All cons items	1.38 (0.31)	1.47 (0.43)	.39
Decisional balance (pros-cons)	1.07 (0.59)	0.82 (0.63)	.18

[a]HPV: human papillomavirus.

The mean vaccination self-efficacy score was higher in the comparison group than in the intervention group (1.65 vs 1.45, respectively; P=.05). Only 3 of the 8 individual items in the scale were significantly different and were in the direction of *confident—not at all*: *When it is too expensive*? *When it is too inconvenient?When my friends will know I got the shot?* The overall mean decisional balance score trended toward greater support of vaccination, although differences between the groups were not significant. As seen in Table 1, *Pros* vaccination scores were higher in the intervention group than in the comparison group (2.44 vs 2.31, respectively), and *cons* vaccination scores were lower in the intervention group than in the comparison group (1.38 vs 1.48, respectively), but again these were not statistically significant differences between the groups.

PENS

The 18 participants who reported on measures of physical/emotional/narrative presence in the game [31] gave mixed reviews on the game. More than half of the participants gave positive scores on game autonomy and competence, ease, and freedom of playing the game (Table 2). At the same time, more than half of the participants called the game boring and said they were not impacted emotionally and that the game did not hold their attention. Thus, the results of this scale revealed both positive and negative evaluations of the game.

Table 2. Measurements from the Physical/Emotional/Narrative Presence Scale (N=18).

Characteristics	Participants who agree or strongly agree, n (%)
In-game autonomy	
The game provides me with interesting options and choices	7 (39)
The game lets you do interesting things	13 (72)
I experienced a lot of freedom in the game	9 (50)
In-game competence	
I feel competent at the game	7 (39)
I feel very capable and effective when playing	8 (44)
My ability to play the game is well-matched with the game's challenges	8 (44)
PENS[a]	
When playing the game, I feel transported to another time and place	4 (22)
Exploring the game world feels like taking an actual trip to a new place	5 (28)
When moving through the game world, I feel as if I am actually there	3 (17)
I am not impacted emotionally by events in the game (−)	9 (50)
The game was emotionally engaging	4 (22)
I experience feelings as deeply in the game as I have in real life	4 (22)
When playing the game, I feel as if I was part of the story	4 (22)
When I accomplished something in the game I experienced genuine pride	6 (33)
I had reactions to events and characters in the game as if they were real	1 (6)
PENS: intuitive controls	
Learning the game controls was easy	10 (56)
The game controls are intuitive	5 (28)
When I wanted to do something in the game, it was easy to remember the corresponding control	10 (56)
Postexperimental Intrinsic Motivation Inventory: interest or enjoyment [23]	
I enjoyed doing this game very much	7 (39)
This game was fun to do	6 (33)
I thought this was a boring game	9 (50)
This game did not hold my attention at all	10 (56)
I would describe this game as very interesting	4 (22)
I thought this game was quite enjoyable	5 (28)
While I was doing this game, I was thinking about how much I enjoyed it	4 (22)
Given the chance I would play this game in my free time	4 (22)
I would like to spend more time playing this game	4 (22)

[a]PENS: Physical/Emotional/Narrative Presence Scale.

Game Play in the Intervention Group

Of the 21 participants assigned to the intervention group, 86% (18) reported playing the game (Table 3).

Preteens who did not play the game reported that they had technical difficulties (n=2), and a parent determined *it was not for me* (n=1). Among the preteens who played the game, 78% (14/18) reported that they played it 3 or more times. The majority of players (14/18, 78%) spent more than 10 min on the game at each session. Although the game was designed primarily for mobile devices (tablets and cellular phones), the majority of the preteens (11/18, 61%) played the game on a personal computer. Half of the participants (9/18, 50%) saved the game to their device. Fifteen of the preteens played the game with a parent, and 4 preteens played with a sibling.

Table 3. Self-reported characteristics of video game play among the intervention group (18 preteens who played the game).

Characteristics	Values, n (%)
Preteen played video game with a sibling	4 (22)
Smartphone or tablet used for video game play	6 (33)
Saved video game to device	9 (50)
Earned at least one shield on the leaderboard	12 (67)
Earned shields on garden plants	8 (44)
Played 3 or more times	14 (78)
Played game more than 10 minutes per session	14 (78)
Preteen played video game with a parent	15 (83)
Played the shield game	13 (72)
Played the hidden objects activity	15 (83)
Created a potion	16 (89)
Correctly identified game metaphor	16 (89)
Playing the game changed how I feel about getting the vaccine	4 (22)
I was interested in finding out more about the HPV[a] vaccine after playing the game	6 (33)
I would recommend this game to a friend who wanted to learn about HPV	6 (33)
I was more confident to talk with a parent about the vaccine after playing the game	8 (44)
I was more willing to get the vaccine after playing the game	9 (50)
I know more about the vaccine	12 (67)

[a]HPV: human papillomavirus.

The participants were asked to complete 3 game activities: (1) play a shield game with blue spikey virus balls, (2) find hidden objects in 4 different rooms, and (3) create a potion (Multimedia Appendix 1). Nearly all the preteens (n=16) found hidden objects and created a potion, and 13 played the shield game. The preteens were able to track their progress on a leaderboard, and 12 of them reported that they earned at least one shield on their leaderboard. The game is completed when shields appear on plants in the garden. Less than half of the preteens (n=8) reported that they saw at least one shield in their garden. All but 2 preteens identified the idea behind the secret garden metaphor.

Postintervention Focus Group

We conducted 3 focus groups with 7 preteens following each of the 3 waves in the intervention. The preteens generally enjoyed and understood the game, especially playing the hidden objects game and earning points to plant their gardens. They acknowledged that playing the game helped them to be more aware of HPV. Participants were curious about what would happen to them if they were vaccinated. They described the game as "just a game where you… just plant flowers in the garden and make a shield to protect the plants," and "...it's a game to help me understand about the HPV shot and what you do in the game." Participants liked the "hidden figures game... were fun to try to find." They said the game was easy to play and that it was fun.

The preteens remembered that messages appeared in the game, but they could not remember specific messages. Study participation did not impact preteens' attitudes about HPV

vaccination, and they agreed that playing the game made them more aware of HPV as an infection. In terms of designing the next level of the video game, they suggested more hidden objects with a higher level of difficulty and a bigger garden as well as pulling weeds out of the garden to take care of the plants. They would include more activities beyond the shield game and the hidden objects.

HPV Vaccination Initiation Rate

The vaccine initiation rate was higher in the intervention group than in the control group, but this difference was not statistically significant (22% vs 15%, respectively; $P=.31$). Vaccination completion rates were also higher in the intervention group than in the control group (9% vs 2%, respectively; $P=.10$). Although this is not significant, it is noteworthy that only 1 of the 27 comparison group members completed the HPV vaccine series, whereas 5 of 28 intervention group members completed the HPV vaccine series. It should also be noted that most of those who initiated were still on schedule to complete, but the date for the next dose had not yet arrived when data were collected.

Discussion

Principal Findings

Of the 36 practices in the study, 18 were able to identify and screen parents and preteens for a total of 55 dyads. The study sample consisted of 21 preteens in the intervention group and 26 preteens in the comparison group who completed the follow-up survey. The objective of our study is to evaluate the acceptability and feasibility of using a serious video game about

HPV vaccination with preteens and parents to promote conversations about and decisions to seek HPV vaccination. The scores for preteen HPV vaccination self-efficacy in our study indicated greater support postintervention for the comparison group compared with the intervention group. Only 3 of the 8 individual items were significant and in the direction of lower self-esteem. One plausible explanation for higher scores in the comparison group is that they were less aware of barriers to HPV vaccination, including expense, inconvenience, and their friends knowing that they would get the vaccine. In addition, game play was mostly positive, with more than half of the participants playing the game as intended and wanting to learn more about HPV vaccination. Less positive comments were made about not changing how they felt about the vaccine or not recommending the game to family or a friend. A greater proportion of preteens in the intervention group initiated the vaccine and had higher completion rates than their counterparts in the control group, but these differences did not reach statistical significance.

From our research and that of others, modifiable determinants to increase HPV vaccination for preteens aged 11 to 12 years include (1) knowledge, attitudes, and beliefs of parents, providers, and preteens; (2) parents' concerns that preteen children are too young to receive vaccination and are not sexually active yet, that the vaccine is not safe, and that they do not have a vaccine recommendation from their doctor; (3) preteens' dislike of shots and being minimally involved in the vaccination decision; and (4) providers' concerns about parents' resistance to vaccination, vaccine cost, and duration [33,34]. Gamification has the potential to increase engagement with health messaging relevant to shaping motivation and behavior, such as seeking HPV vaccination [21]. Gamification includes techniques to increase knowledge and shape attitudes about HPV vaccination. These techniques often provoke positive effects, depending on how they are being implemented and used [21]. The use of a garden metaphor, for example, in visualizing the importance of HPV vaccination, facilitates the preteen's conception of a beneficial medical procedure to prevent harmful viruses. In the case of the *Land of Secret Gardens*, the challenge is to grow a healthy garden, protected from viruses.

The strengths of our study include using a serious video game to motivate interest in HPV vaccination and to promote conversations with parents, family members, and friends. We conducted focus groups with preteens and learned their viewpoints about serious video games. We further conducted focus groups with preteens as we built the game to determine functionality. Finally, we conducted focus groups after game play to learn what worked well and what did not work so well. Our thorough process will help make the game more relevant to the preteens.

Limitations

The small sample size was a primary weakness of this study. Recruitment and retention were barriers throughout the study. Once clinic staff provided names of potential participants to research staff, follow-up with the parents via our protocol of 3 attempts via both email and telephone contact remained to be a challenge. One obstacle was a high proportion of clients with parents needing informed consent provided in Spanish, an accommodation not available for this study. Another obstacle reflected some of the measurements we had available. For instance, the decisional balance measure was developed with older subjects (college-aged women); therefore, some of the items had to be modified for use in a young population of children aged 11 to 12 years. A further limitation was relying on vaccination and self-efficacy results from a group of older teenagers. This might have skewed the results from younger teenagers.

Conclusions

A serious video game on HPV vaccination is acceptable to parents and preteens and can be played as intended. Gamification can be effective in shaping attitudes about the HPV vaccination. Further research is needed to enhance the game with puzzles and activities that are engaging to the preteen population.

Authors' Contributions

JLC, BFF, SJD, and JRC designed the study. BFF, LLS, SJD, JRC, and TCB administered the surveys or conducted the interviews. JRC, BFF, LLS, SJD, and JLC analyzed the study data. JLC, BFF, LLS, SJD, JRC, and TCB wrote the manuscript.

Conflicts of Interest

At the time of this study TCB served on an advisory board for Pfizer Inc. All other authors declare no conflicts.

Editorial Notice

This randomized study was only retrospectively registered; the study was initiated by the University of North Carolina at Chapel Hill prior to this requirement being requested by the journal. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and the study was considered formative, guiding the development of the application. However, readers are advised to carefully assess the validity of any potential explicit or

implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
Video game for preteens.
[PPTX File , 7094 KB - games_v8i4e16883_app1.pptx]

Multimedia Appendix 2
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File , 344 KB - games_v8i4e16883_fig.pdf]

References

1. Walker TY, Elam-Evans LD, Yankey D, Markowitz LE, Williams CL, Fredua B, et al. National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years - United States, 2018. MMWR Morb Mortal Wkly Rep 2019 Aug 23;68(33):718-723 [FREE Full text] [doi: 10.15585/mmwr.mm6833a2] [Medline: 31437143]

2. Walker TY, Elam-Evans LD, Yankey D, Markowitz LE, Williams CL, Mbaeyi SA, et al. National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years - United States, 2017. MMWR Morb Mortal Wkly Rep 2018 Aug 24;67(33):909-917 [FREE Full text] [doi: 10.15585/mmwr.mm6733a1] [Medline: 30138305]

3. Vielot NA, Butler AM, Brookhart MA, Becker-Dreps S, Smith JS. Patterns of use of human papillomavirus and other adolescent vaccines in the United States. J Adolesc Health 2017 Sep;61(3):281-287. [doi: 10.1016/j.jadohealth.2017.05.016] [Medline: 28739327]

4. Holman DM, Benard V, Roland KB, Watson M, Liddon N, Stokley S. Barriers to human papillomavirus vaccination among US adolescents: a systematic review of the literature. JAMA Pediatr 2014 Jan;168(1):76-82 [FREE Full text] [doi: 10.1001/jamapediatrics.2013.2752] [Medline: 24276343]

5. Thompson EL, Rosen BL, Vamos CA, Kadono M, Daley EM. Human papillomavirus vaccination: what are the reasons for nonvaccination among US adolescents? J Adolesc Health 2017 Sep;61(3):288-293. [doi: 10.1016/j.jadohealth.2017.05.015] [Medline: 28842066]

6. Niccolai LM, Hansen CE. Practice- and community-based interventions to increase human papillomavirus vaccine coverage: a systematic review. JAMA Pediatr 2015 Jul;169(7):686-692 [FREE Full text] [doi: 10.1001/jamapediatrics.2015.0310] [Medline: 26010507]

7. Smulian EA, Mitchell KR, Stokley S. Interventions to increase HPV vaccination coverage: a systematic review. Hum Vaccin Immunother 2016 Jun 2;12(6):1566-1588 [FREE Full text] [doi: 10.1080/21645515.2015.1125055] [Medline: 26838959]

8. Mercer SL, DeVinney BJ, Fine LJ, Green LW, Dougherty D. Study designs for effectiveness and translation research :identifying trade-offs. Am J Prev Med 2007 Aug;33(2):139-154. [doi: 10.1016/j.amepre.2007.04.005] [Medline: 17673103]

9. Dempsey AF, Pyrznawoski J, Lockhart S, Barnard J, Campagna EJ, Garrett K, et al. Effect of a health care professional communication training intervention on adolescent human papillomavirus vaccination: a cluster randomized clinical trial. JAMA Pediatr 2018 May 7;172(5):e180016 [FREE Full text] [doi: 10.1001/jamapediatrics.2018.0016] [Medline: 29507952]

10. Markowitz LE, Gee J, Chesson H, Stokley S. Ten years of human papillomavirus vaccination in the United States. Acad Pediatr 2018 Mar;18(2S):S3-10. [doi: 10.1016/j.acap.2017.09.014] [Medline: 29502635]

11. Cates JR, Crandell JL, Diehl SJ, Coyne-Beasley T. Immunization effects of a communication intervention to promote preteen HPV vaccination in primary care practices. Vaccine 2018 Jan 2;36(1):122-127 [FREE Full text] [doi: 10.1016/j.vaccine.2017.11.025] [Medline: 29169892]

12. Cates JR, Diehl SJ, Crandell JL, Coyne-Beasley T. Intervention effects from a social marketing campaign to promote HPV vaccination in preteen boys. Vaccine 2014 Jul 16;32(33):4171-4178 [FREE Full text] [doi: 10.1016/j.vaccine.2014.05.044] [Medline: 24886960]

13. Cates JR, Shafer A, Diehl SJ, Deal AM. Evaluating a county-sponsored social marketing campaign to increase mothers' initiation of HPV vaccine for their pre-teen daughters in a primarily rural area. Soc Mar Q 2011;17(1):4-26 [FREE Full text] [doi: 10.1080/15245004.2010.546943] [Medline: 21804767]

14. Cates JR, Ortiz R, Shafer A, Romocki LS, Coyne-Beasley T. Designing messages to motivate parents to get their preteenage sons vaccinated against human papillomavirus. Perspect Sex Reprod Health 2012 Mar;44(1):39-47 [FREE Full text] [doi: 10.1363/4403912] [Medline: 22405151]

15. Shafer A, Cates JR, Diehl SJ, Hartmann M. Asking mom: formative research for an HPV vaccine campaign targeting mothers of adolescent girls. J Health Commun 2011 Oct;16(9):988-1005. [doi: 10.1080/10810730.2011.571343] [Medline: 21728780]

16. Cates JR, Diehl SJ, Fuemmeler BF, North SW, Chung RJ, Hill JF, et al. Toward optimal communication about HPV vaccination for preteens and their parents: evaluation of an online training for pediatric and family medicine health care providers. J Public Health Manag Pract 2020;26(2):159-167. [doi: 10.1097/PHH.0000000000001022] [Medline: 31348153]

17. Patchen L, Ellis L, Ma TX, Ott C, Chang KH, Araya B, et al. Engaging African American youth in the development of a serious mobile game for sexual health education: mixed methods study. JMIR Serious Games 2020 Jan 31;8(1):e16254 [FREE Full text] [doi: 10.2196/16254] [Medline: 32012041]

18. Eley CV, Young VL, Hayes CV, Verlander NQ, McNulty CA. Young people's knowledge of antibiotics and vaccinations and increasing this knowledge through gaming: mixed-methods study using e-bug. JMIR Serious Games 2019 Feb 1;7(1):e10915 [FREE Full text] [doi: 10.2196/10915] [Medline: 30707096]

19. Thompson D, Baranowski T, Buday R, Baranowski J, Thompson V, Jago R, et al. Serious video games for health how behavioral science guided the development of a serious video game. Simul Gaming 2010 Aug 1;41(4):587-606 [FREE Full text] [doi: 10.1177/1046878108328087] [Medline: 20711522]

20. Lu AS, Baranowski J, Islam N, Baranowski T. How to engage children in self-administered dietary assessment programmes. J Hum Nutr Diet 2014 Jan;27(Suppl 1):5-9 [FREE Full text] [doi: 10.1111/j.1365-277X.2012.01258.x] [Medline: 22594618]

21. Baranowski T, Buday R, Thompson DI, Baranowski J. Playing for real: video games and stories for health-related behavior change. Am J Prev Med 2008 Jan;34(1):74-82 [FREE Full text] [doi: 10.1016/j.amepre.2007.09.027] [Medline: 18083454]

22. Thompson D, Baranowski T, Buday R. Conceptual model for the design of a serious video game promoting self-management among youth with type 1 diabetes. J Diabetes Sci Technol 2010 May 1;4(3):744-749 [FREE Full text] [doi: 10.1177/193229681000400331] [Medline: 20513343]

23. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. Motiv Emot 2006 Nov 29;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]

24. Cates JR, Fuemmeler BF, Diehl SJ, Stockton LL, Porter J, Ihekweazu C, et al. Developing a serious videogame for preteens to motivate HPV vaccination decision making: land of secret gardens. Games Health J 2018 Feb;7(1):51-66 [FREE Full text] [doi: 10.1089/g4h.2017.0002] [Medline: 29161529]

25. Meites E, Kempe A, Markowitz LE. Use of a 2-dose schedule for human papillomavirus vaccination - updated recommendations of the advisory committee on immunization practices. MMWR Morb Mortal Wkly Rep 2016 Dec 16;65(49):1405-1408 [FREE Full text] [doi: 10.15585/mmwr.mm6549a5] [Medline: 27977643]

26. Gamification. Merriam Webster Inc. 2019. URL: https://www.merriam-webster.com/dictionary/gamification/ [accessed 2020-09-16]

27. Ryan D. Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp Educ Psychol 2000 Jan;25(1):54-67. [doi: 10.1006/ceps.1999.1020] [Medline: 10620381]

28. McGonigal J. Reality is Broken : Why Games Make Us Better and How They Can Change the World. New York, USA: Penguin Press; 2011.

29. Survey Research Suite. Qualtrics Inc. 2018. URL: https://www.qualtrics.com/blog/citing-qualtrics/ [accessed 2020-09-16]

30. Navicat Database Management. 2018. URL: https://navicat.com/en/ [accessed 2018-10-07]

31. Busselle R, Bilandzic H. Measuring narrative engagement. Media Psychol 2009 Nov 23;12(4):321-347. [doi: 10.1080/15213260903287259]

32. Lipschitz JM, Fernandez AC, Larson HE, Blaney CL, Meier KS, Redding CA, et al. Validation of decisional balance and self-efficacy measures for HPV vaccination in college women. Am J Health Promot 2013;27(5):299-307. [doi: 10.4278/ajhp.110606-QUAN-240] [Medline: 23402229]

33. National and State Vaccination Coverage Among Adolescents Aged 13–17 Years — United States, 2011. Centers for Disease Control and Prevention. 2013. URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6134a3.htm?s_cid=mm6134a3_w [accessed 2019-09-04]

34. Brewer NT, Gottlieb SL, Reiter PL, McRee AL, Liddon N, Markowitz L, et al. Longitudinal predictors of human papillomavirus vaccine initiation among adolescent girls in a high-risk geographic area. Sex Transm Dis 2011 Mar;38(3):197-204 [FREE Full text] [doi: 10.1097/OLQ.0b013e3181f12dbf] [Medline: 20838362]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act
HPV: human papillomavirus
IRB: institutional review board
PENS: Physical/Emotional/Narrative Presence Scale
Tdap: tetanus, diphtheria, and acellular pertussis

XSL•FO

RenderX

XSL•FO
RenderX

Original Paper

A Serious Game Designed to Promote Safe Behaviors Among Health Care Workers During the COVID-19 Pandemic: Development of "Escape COVID-19"

Mélanie Suppan[1], MD; Gaud Catho[2], MD; Tomás Robalo Nunes[2,3], MD; Valérie Sauvan[2], RN; Monique Perez[2], RN; Christophe Graf[4], MD; Didier Pittet[2], MD, MS; Stephan Harbarth[2], MD, MSc; Mohamed Abbas[2], MD; Laurent Suppan[5], MD

[1]Division of Anesthesiology, Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, University of Geneva Hospitals and Faculty of Medicine, Geneva, Switzerland

[2]Infection Control Program, WHO Collaborating Centre on Patient Safety, University of Geneva Hospitals and Faculty of Medicine, Geneva, Switzerland

[3]Infectious Diseases Department, Hospital Garcia de Orta, EPE, Almada, Portugal

[4]Department of Rehabilitation and Geriatrics, Geneva University Hospital, Geneva, Switzerland

[5]Division of Emergency Medicine, Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, University of Geneva Hospitals and Faculty of Medicine, Geneva, Switzerland

Corresponding Author:
Mélanie Suppan, MD
Division of Anesthesiology
Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine
University of Geneva Hospitals and Faculty of Medicine
Rue Gabrielle-Perret-Gentil 4
Geneva, 1211
Switzerland
Phone: 41 223723311
Email: melanie.bochet@hcuge.ch

Abstract

Background: As many countries fear and even experience the emergence of a second wave of COVID-19, reminding health care workers (HCWs) and other hospital employees of the critical role they play in preventing SARS-CoV-2 transmission is more important than ever. Building and strengthening the intrinsic motivation of HCWs to apply infection prevention and control (IPC) guidelines to avoid contaminating their colleagues, patients, friends, and relatives is a goal that must be energetically pursued. A high rate of nosocomial infections during the first COVID-19 wave was detected by IPC specialists and further cemented their belief in the need for an engaging intervention that could improve compliance with COVID-19 safe behaviors.

Objective: Our aim was to develop a serious game that would promote IPC practices with a specific focus on COVID-19 among HCWs and other hospital employees.

Methods: The first 3 stages of the SERES framework were used to develop this serious game. A brainswarming session between developers and IPC specialists was used to identify the target audience and acquisition objectives. Nicholson's RECIPE mnemonic (reflection, engagement, choice, information, play, exposition) for meaningful gamification was used to guide the general design. A common and simple terminology was used to suit the broad target audience. The game was tested on various platforms (smartphones, tablets, laptops, desktop computers) by different users during each development loop and before its final release.

Results: The game was designed to target all hospital staff who could be in direct contact with patients within the Geneva University Hospitals. In total, 10 acquisition objectives were defined by IPC specialists and implemented into the game according to the principles of meaningful gamification. A simple storyboard was first created using Microsoft PowerPoint and was progressively refined through multiple iteration loops. Articulate Storyline was then used to create two successive versions of the actual game. In the final version, a unique graphic atmosphere was created with help from a professional graphic designer. Feedback mechanisms were used extensively throughout the game to strengthen key IPC messages.

Conclusions: The SERES framework was successfully used to create "Escape COVID-19," a serious game designed to promote safe IPC practices among HCWs and other hospital employees during the COVID-19 pandemic. This game can be obtained free

XSL•FO
RenderX

of charge for research and educational purposes. A SCORM (shareable content object reference model) package is available to facilitate results and completion tracking on most current learning management systems.

Introduction

Background and Importance

As the COVID-19 death toll continues to rise [1] and many countries fear or are experiencing the emergence of a second wave [2-4], avoiding disease transmission is becoming increasingly important. Health care workers (HCWs) and other hospital employees, including medical, nursing, and administrative and support staff, are at high risk of contamination [5-7]. The actual rate of COVID-19 infection among HCWs is unknown and may vary significantly from one country to another. Recent studies have shown that as many as 20% of asymptomatic HCWs had serological evidence of having been exposed to SARS-CoV-2, the virus responsible for COVID-19 [8]. Inadequate application of infection prevention and control (IPC) procedures by HCWs and other hospital employees can lead to the contamination of their colleagues, patients, and even visitors [9]. Improper handling of personal protective equipment (PPE) can also lead to self-contamination [10,11].

As the pandemic draws out, lassitude and conspiracy theories can lead to a slackening of safety measures [12,13], including among HCWs. Moreover, some HCWs seem to lack critical knowledge regarding the modes of transmission of SARS-CoV-2 [14]. In addition, the frequent and sometimes conflicting updates pertaining to COVID-19 guidelines have led many HCWs to question their applicability and protective value [15]. This can lead to a lack of commitment, one of the barriers that negatively impacts the implementation of measures necessary to prevent SARS-CoV-2 transmission [16]. Regular channels used to spread IPC messages might therefore fail in the current context.

Nevertheless, building and strengthening the intrinsic motivation of HCWs to rigorously apply IPC guidelines and to do their very best to avoid contaminating their patients, colleagues, friends, and relatives is a goal that must be energetically pursued [17]. Apart from the significant morbidity and mortality associated with COVID-19 infection [18], the economic burden is also worth considering. Indeed, the median costs incurred for a single case exceed $3000 according to a recent evaluation [19].

The high rate of nosocomial infections detected during the first COVID-19 wave and the need to train a large number of people while respecting the need for physical distance require the creation of original and inspirational material to promote IPC guidelines. Serious games possess motivational properties that can enhance learner engagement and satisfaction while still delivering important messages [20].

Objective

Our aim was to develop a serious game to promote the dissemination of COVID-19 IPC practices among HCWs and other hospital employees. By raising awareness about situations that could potentially lead to COVID-19 contamination in this key population, such a game could help decrease SARS-CoV-2 transmission between HCWs and from HCWs to patients.

Methods

General Design

The first 3 stages of the SERES framework were used to develop this serious game [21]. This framework has already been used successfully for the development of serious games [22] and gamified e-learning (electronic learning) modules [23]. Briefly, this framework was created to help developers design theory-driven, evidence-based serious games for health. The first stage, scientific foundations, is designed to ensure that the game is both evidence-based and theoretically driven. The target audience is identified at this stage. The second stage, design foundations, focuses on the translation of these foundations into design elements. Game development takes place in the third stage of this framework.

Scientific Foundations

Target Audience

The target audience was identified through a brainswarming session [24] between the lead developer (MS) and IPC specialists (GC, TRN, VS, MP, SH, and MA). The game was developed at the request of these specialists, who were therefore included in the development process from the start.

Acquisition Objectives

Field observations by specialists from the Geneva University Hospitals infection control program were used to identify the acquisition objectives. We decided to use the term "acquisition objectives" rather than "learning objectives" as the main purpose of this serious game was to promote and stimulate desirable behaviors in HCWs rather than to teach them new procedures.

Theoretical Basis

The theoretical bases relevant to the creation of gamified content in the context of the COVID-19 pandemic were already identified in the course of a previously published project [23]. Therefore, we decided to use these bases and to reassess the relevant references. Briefly, we had searched the medical literature via PubMed (using Medical Subject Headings [MeSH] and Boolean operators) and retrieved articles based on their titles and abstracts. References from the most relevant articles were also searched manually to identify sources that could have

been missed. A total of 12 articles was included through this procedure. Further details are available in our previous publication [23].

We also performed a search for articles reporting information related to SARS-CoV-2 transmission. The MeSH keywords "covid-19," "sars-cov-2," and "transmission" were used in combination with the Boolean operators "AND" and "OR" to retrieve potentially relevant references from MEDLINE using the PubMed engine. To avoid inducing doubts or even mistrust in HCWs [25-27], the results retrieved through this search were anaylzed to determine whether they were in line with our institutional [28] and federal [29] guidelines. Whenever conflicting information was identified, a consensus was reached with the IPC specialists to determine and use the most accurate and least confusing messages.

Content Validation

The scientific grounds on which this serious game relies were decided during brainswarming sessions between the developers and IPC specialists from the Geneva University Hospitals. Each iteration of the serious game was reviewed by at least 2 such specialists to ensure that the intended message was correctly and adequately conveyed by the serious game.

Design Foundations

General Design and Meaningful Gamification

The scientific foundations reviewed or established during the previous stage were used to create the general design of the serious game. We used Nicholson's RECIPE mnemonic (reflection, engagement, choice, information, play, and exposition) for meaningful gamification to guide the development of the game mechanics [30].

Arnab et al's [31] model of transforming learning mechanics into game mechanics was adapted to suit the development of this serious game. The original model states that learning mechanics should be decided according to learning objectives and that these learning mechanics should then be translated into game mechanics. As we had decided to use the term "acquisition objectives" rather than "learning objectives," we adapted Arnab et al's [31] model to develop a very similar "acquisition mechanics – game mechanics" model.

Design Requirements

Design decisions were tailored to the target population and the COVID-19 context. Ease of access to the serious game was considered to be of paramount importance, as was the ability to rapidly capture the player's attention and interest. True to the SERES framework, we decided to use an iterative approach to build the game. We therefore sought regular feedback from the IPC specialists as well as from potential end users and implemented the required adaptations accordingly.

Game Development

All the data gathered through the scientific foundation and design foundation stages were used to choose a development platform and to proceed with the development of the serious game. Acquisition objectives were reassessed, and theoretical

bases were used to decide upon acquisition mechanics and game mechanics.

Tool Evaluation

In line with the theoretical bases, the game was tested on various platforms (smartphones, tablets, laptops, desktop computers) by different users at the end of each development loop and before its final release.

Results

Scientific Foundations

Target Audience

Early on in the development process, developers and IPC specialists agreed that the target audience should be composed of all the professionals who could be in direct contact with patients within the Geneva University Hospitals. The game was therefore designed to target many different kinds of health care professionals including physicians, nurses, physiotherapists, assistants, technicians, and therapists but also non–health care professionals. Indeed, some of these employees have close and regular contact with patients (stretcher-bearers, housekeepers, etc) and might therefore be exposed to patients or play a role in SARS-CoV-2 transmission.

Theoretical Basis

We used previously identified resources to establish the theoretical background [20,31-37]. These resources were extracted from Verschueren et al [22], which described the development of a serious game designed to reduce perioperative anxiety in children, and from our recent study describing the development of a gamified e-learning module that was used to teach the adequate use of PPE to prehospital personnel in the context of the COVID-19 pandemic [23]. Moreover, we sought to include all 6 elements of Nicholson's meaningful gamification [30]. These elements are listed below and follow Nicholson's RECIPE mnemonic:

- The "reflection" element seeks to connect the game to events that happen or might happen to the player in real life. This element was therefore quite straightforward to embed into the game as it was linked to its core concept. We decided to make the player virtually experience scenes that could occur on a normal day to let them make choices that could influence SARS-CoV-2 transmission. We decided to split the game into 4 levels, each related to a specific part of the day (Figures 1 and 2): at home, on the way to work, in communal areas, and in the ward. Including the reflection element has been shown to enhance mental well-being [17] and could thus increase the player's willingness to adopt desirable behaviors.
- There are two components to the "engagement" element. Social engagement such as that found in multiplayer games is the component that has gained the most popularity in recent history. While such games have shown many positive features, they are also controversial and have been accused of drawing some players away from reality [38]. Given the limited time frame and the goal of our serious game, the inclusion of multiplayer capabilities was not considered

relevant. We nevertheless decided to include a social engagement component in the game and to confront the player with the impact some of their choices would have on others. This component was also instrumental in strengthening the reflection element and in keeping the player connected to the real world. The second component of the engagement element is related to the creation of an engaging learning experience and to the concept of "flow" [39]. To avoid boredom and to increase engagement, the challenges should become more difficult as the game progresses. Therefore, simple questions (Figures 3-5) were first asked at the initial level ("at home"), while more complex challenges, such as rebuilding a donning sequence (Figure 6) or choosing the kind of face mask to wear (Figure 7) awaited the player in the last level ("ward"). The main drawback of creating more complex challenges is the risk of generating anxiety and frustration. To avoid these negative effects, feedback mechanisms were used extensively [32].

- The "choice" element relates to the autonomy the player has within the game. Giving the player the ability to make meaningful choices reinforces the player's autonomy and the feeling of being responsible for their actions. We therefore decided to let players make choices that IPC specialists would disapprove, and to experience (at least virtually) the consequences such potentially harmful choices might have.

- "Information" means providing the player with key concepts to help them understand the reasons behind the serious game, rather than just providing them with points or rewards, which could prove ineffective or even harmful in

the long run [40]. Giving players accurate and reliable information ultimately leads to the appropriation of key concepts and generates a positive mental image by generating a feeling of "mastery" regarding these concepts [17]. We therefore included the most recent and evidence-based elements regarding SARS-CoV-2 symptoms and transmission within the game [41-46].

- The "play" element is perhaps the hardest to define, and there are many and sometimes conflicting definitions of this element. We have elected to use Nicholson's approach [30] in which play is defined as "the freedom to explore and fail within boundaries." We therefore decided to let players have a certain degree of freedom and to make choices, which may ultimately result in a "game over" situation. The player would then be given the opportunity either to restart the level or to trade tokens to resume the game.

- Exposition means creating a meaningful narrative in the serious game. Engaging the player into the serious game requires capturing their interest and curiosity while still permitting them to make their own choices. To create such a narrative, we decided that the player would virtually go through the steps they usually undertake on a daily basis, and to make choices that may lead to the contamination of patients or of fellow HCWs.

Audio elements were not considered necessary as many hospital employees will play the serious game on institutional computers and will not be able to activate the sound. Attractive graphics adapted to the target population were however deemed essential to increase engagement [47,48].

Figure 1. First version of the welcome screen.

En pleine pandémie de COVID-19 les règles d'hygiène ne sont plus respectées et le virus se répend dans la population. En tant qu'employé des HUG, **vous jouez un rôle essentiel pour stopper les chaînes de transmission.** Cliquez sur l'un des niveaux pour commencer. Vous devrez compléter chaque niveau pour terminer cette formation.
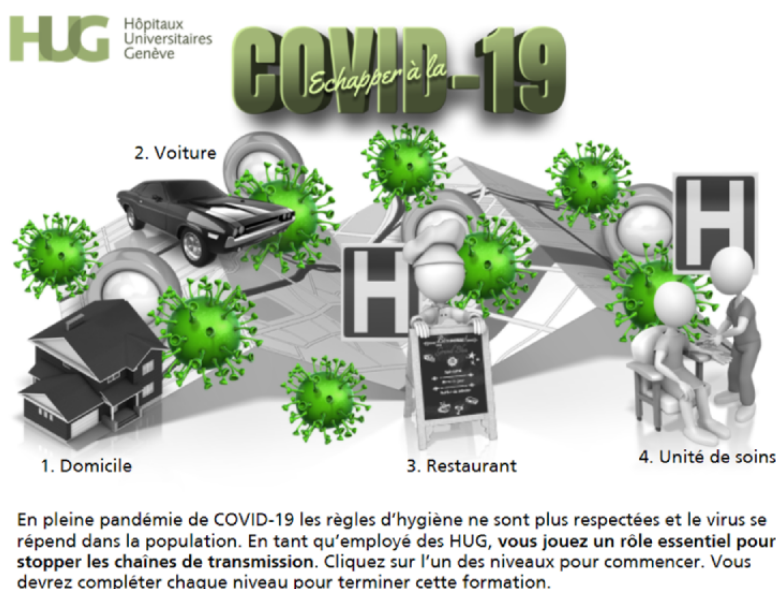
Figure 2. Final version of the welcome screen.
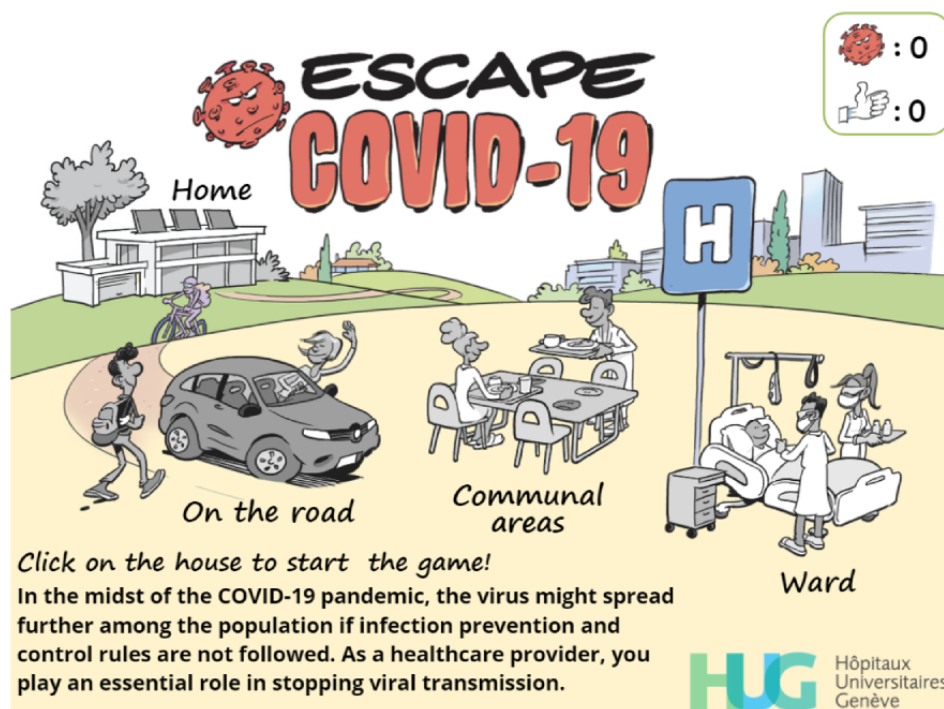
Figure 3. First version of a simple-choice interaction.
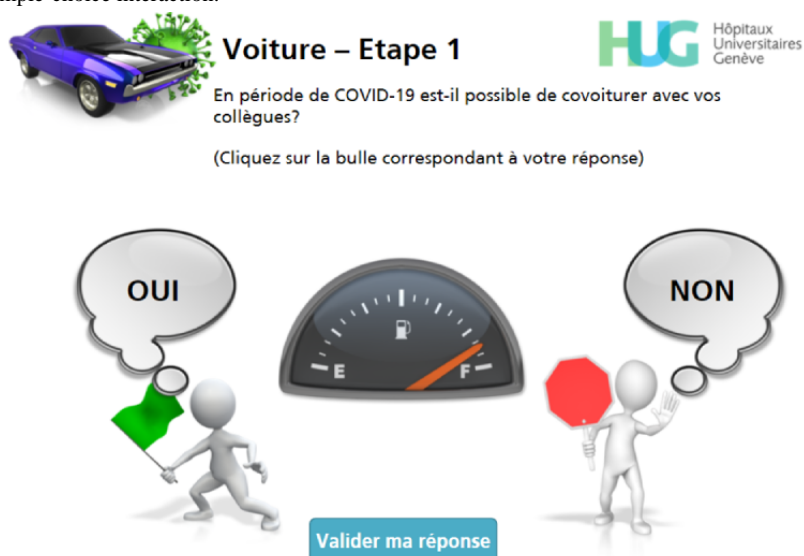
Figure 4. Final version of a simple-choice interaction.
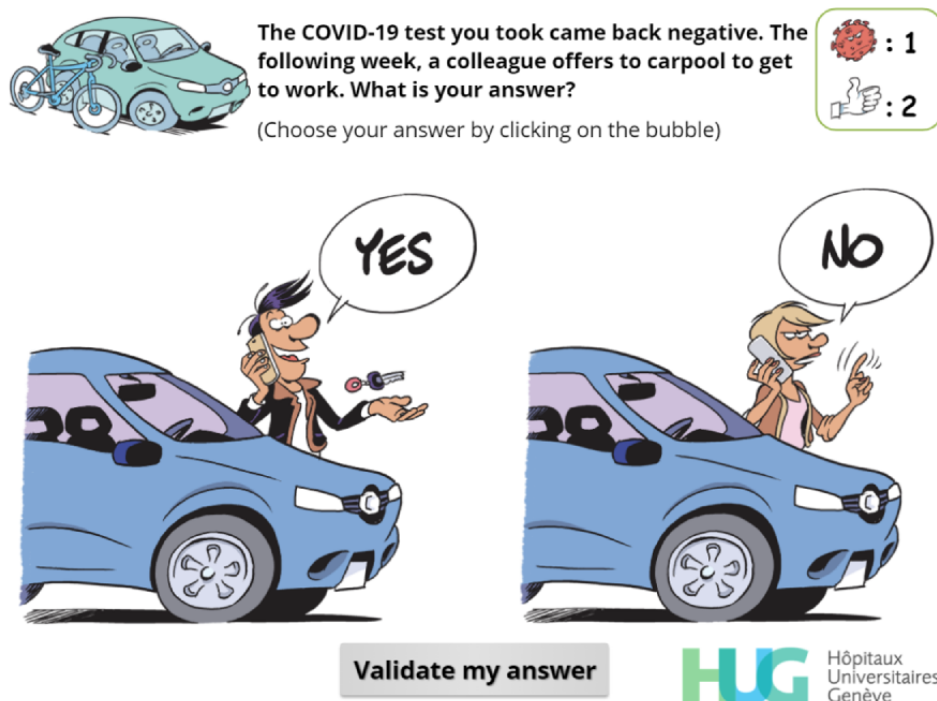
Figure 5. Simple-choice interaction.
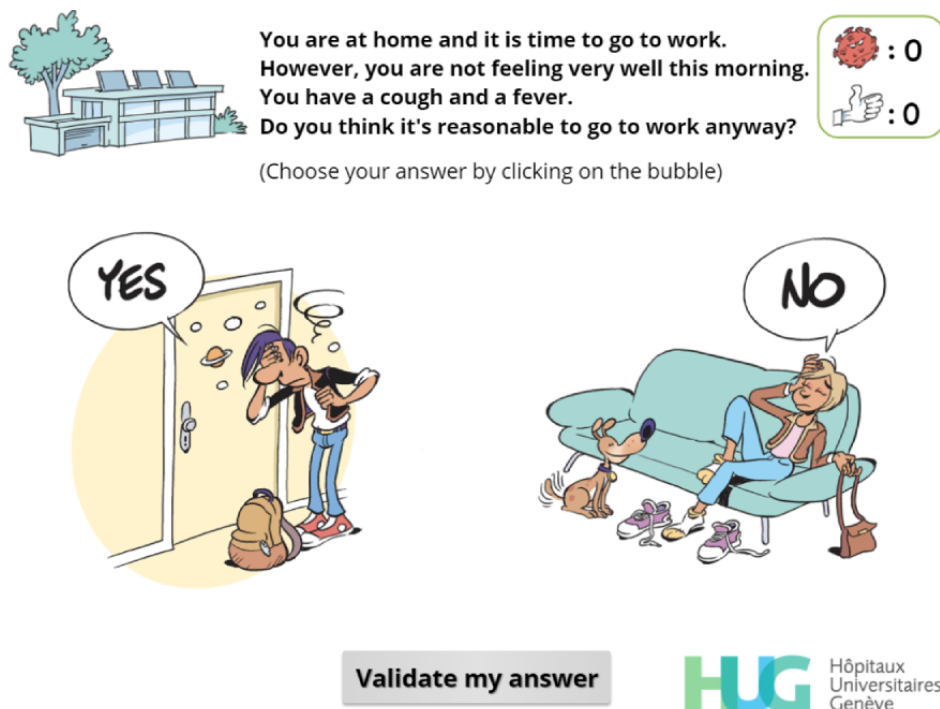
Figure 6. Rebuilding the donning sequence. The player must drag and drop the elements of personal protective equipment in the right order.
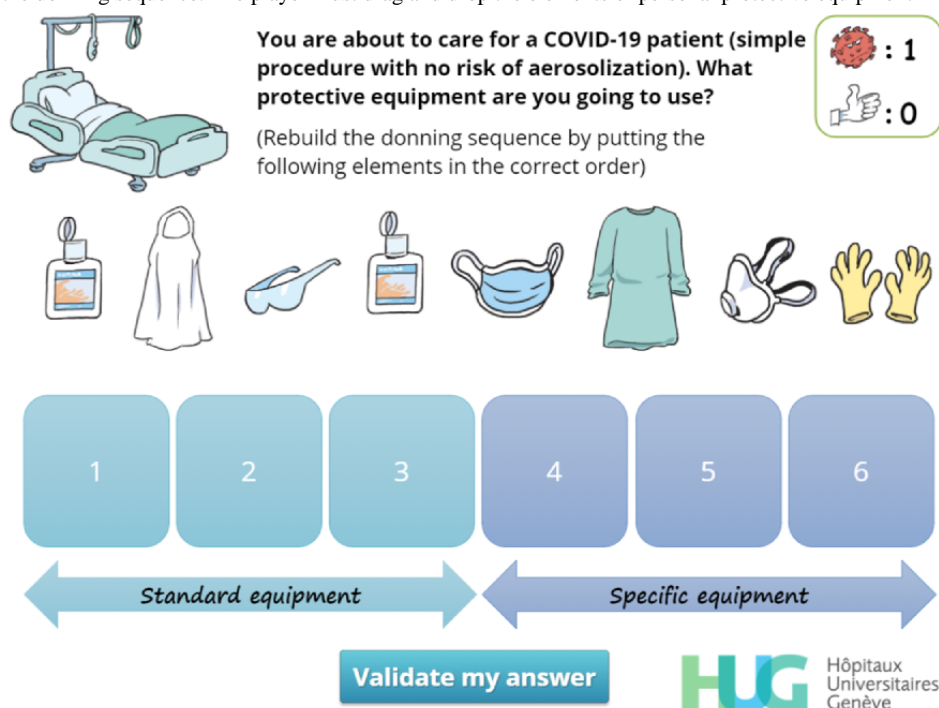
Figure 7. Multiple-choice interaction. The player must choose in which situations N95 masks should be worn. Illustrations change as the cursor hovers over the buttons on the left.

Acquisition Objectives

The IPC specialists defined 10 acquisition objectives according to their field observations. Table 1 summarizes these objectives, the RECIPE element to which they are linked, and gives implementation examples.

Table 1. Acquisition objectives, RECIPE (R=reflection, EN=engagement, C=choice, I=information, P=play, and EX=exposition) elements, and implementation examples.

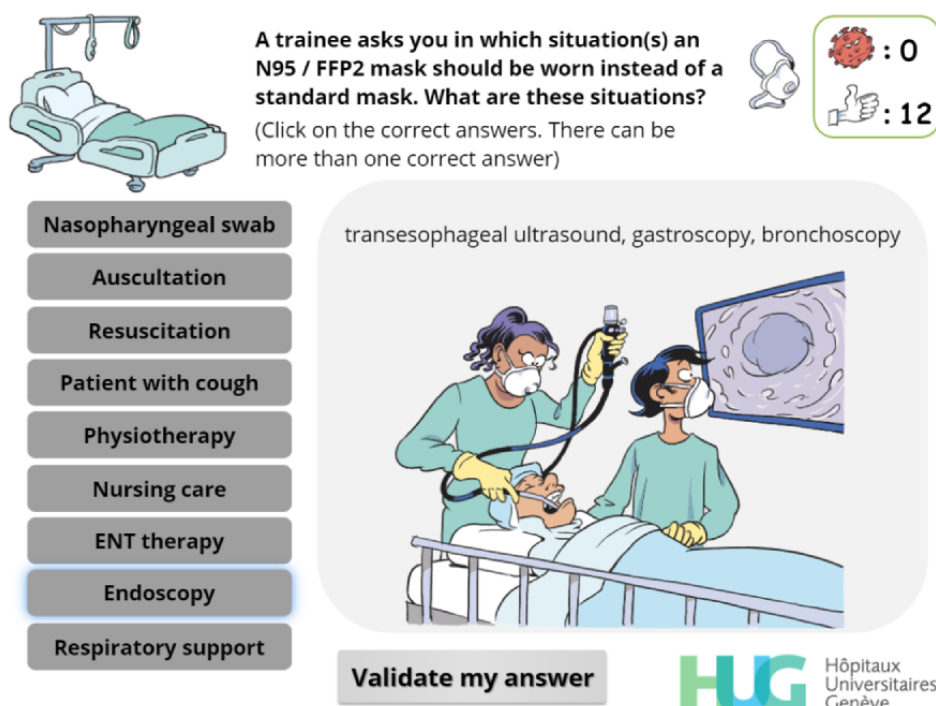| Acquisition objective	RECIPE elements	Implementation
Avoid going to work when potentially infectious	R, EN, C, P, EX	Figures 5,8, 9
Recognize COVID-19 symptoms	I	Figure 10
Recognize that HCWs[a] have a higher probability of being contaminated by a colleague than by a patient	R, EN, I, EX	Figure 11
Recognize that even asymptomatic carriers can be contagious	EN, I, EX	Figure 12
Recall safety measures in the working environment	R, I, EX	Figures 13 and 14
Recall the correct way of using a face mask	R, I, EX	Figure 15
Recall the correct donning sequence	R, EN, I, EX	Figure 6
Identify when an N95 mask should be worn	R, C, I, EX	Figure 7
Identify when eye protection should be worn	R, C, I, EX	Figure 16
Identify when gloves should be worn and changed	R, C, I, EX	Figure 17

[a]HCW: health care worker.

Figure 8. Feedback. The player has chosen to go to work despite the symptoms and is now told they should have undergone a test to screen for SARS-CoV-2 infection. The virus count (top right) has increased accordingly.
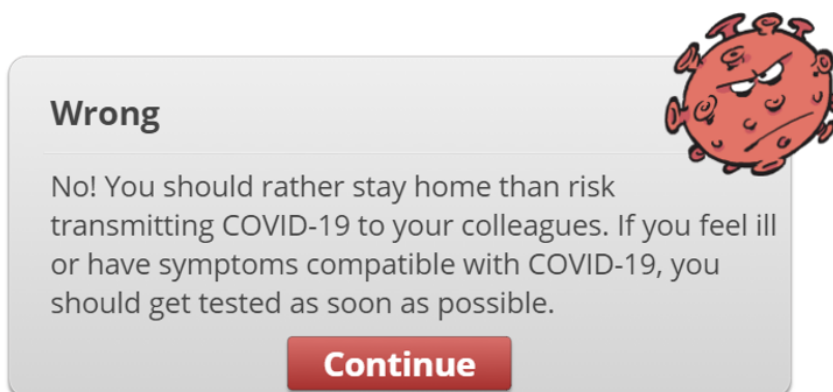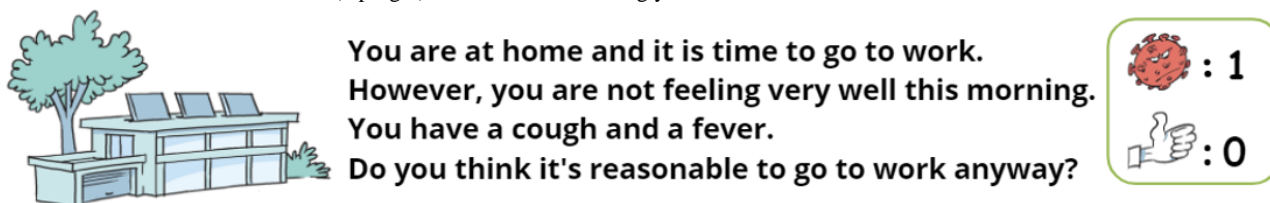
Figure 9. Feedback. The player has chosen not to go to work, and this choice is encouraged by a positive message and by a "thumbs-up" image. The "thumbs-up" count (top right) has increased accordingly.
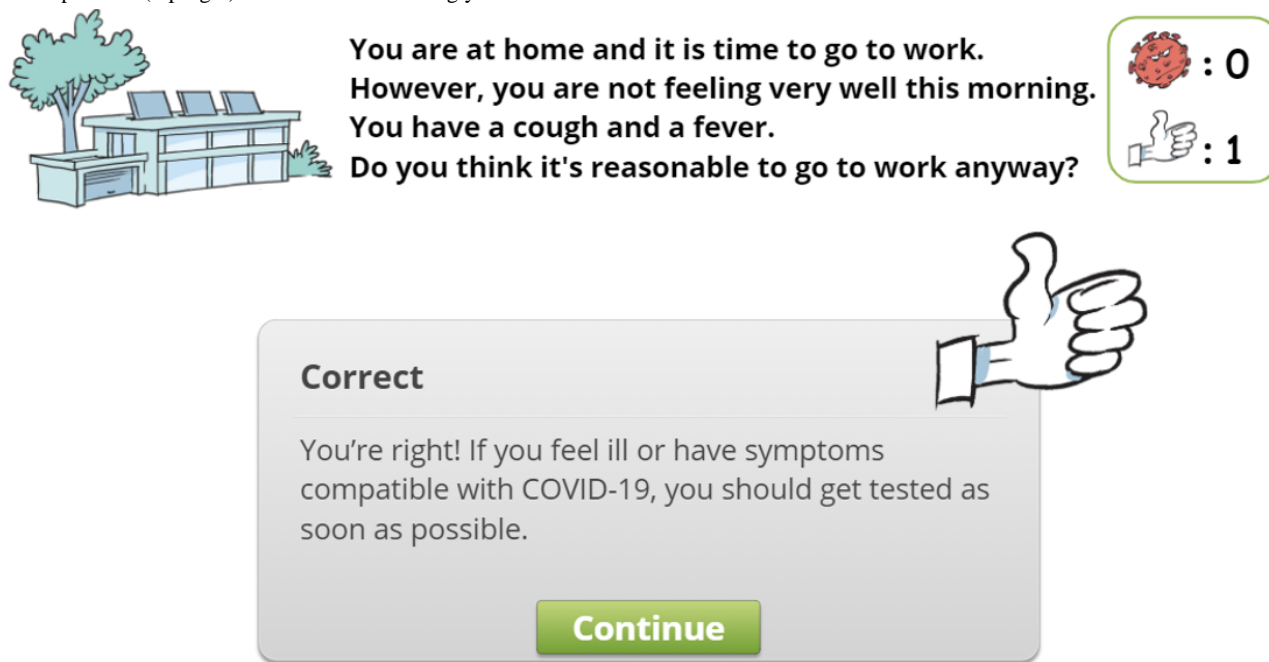
Figure 10. Multiple-choice interaction. The player has to identify the symptoms that should prompt a COVID-19 screening test.

XSL•FO

RenderX
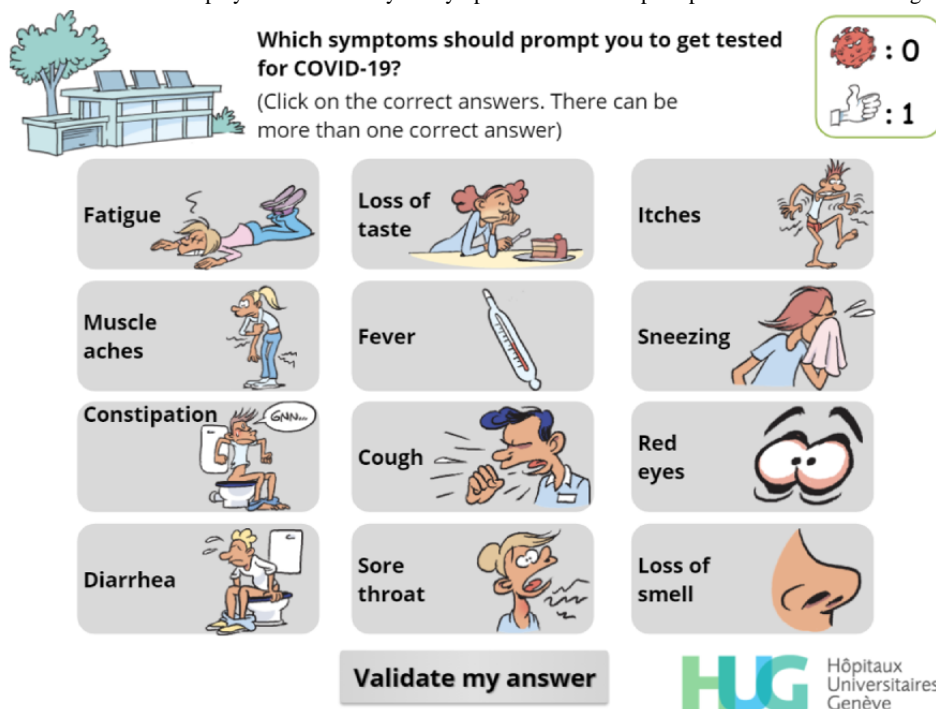
Figure 11. Simple-choice interaction. The player has not selected any option and cannot click on "Validate my answer" for the moment.
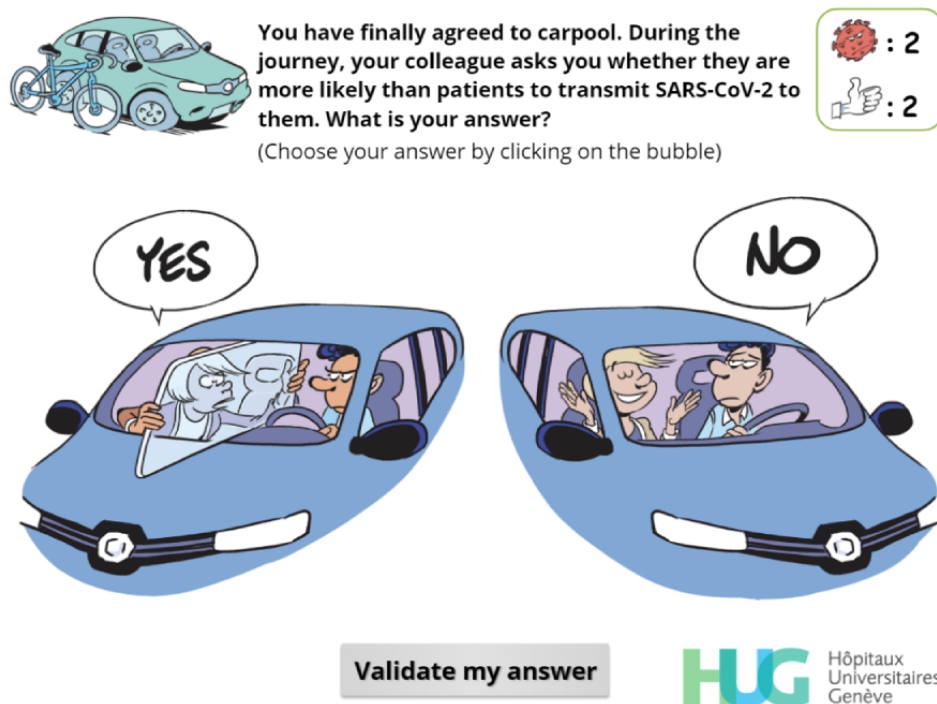
Figure 12. Simple-choice interaction. The player has already clicked on "Yes" and the "Validate my answer" button can now be clicked.
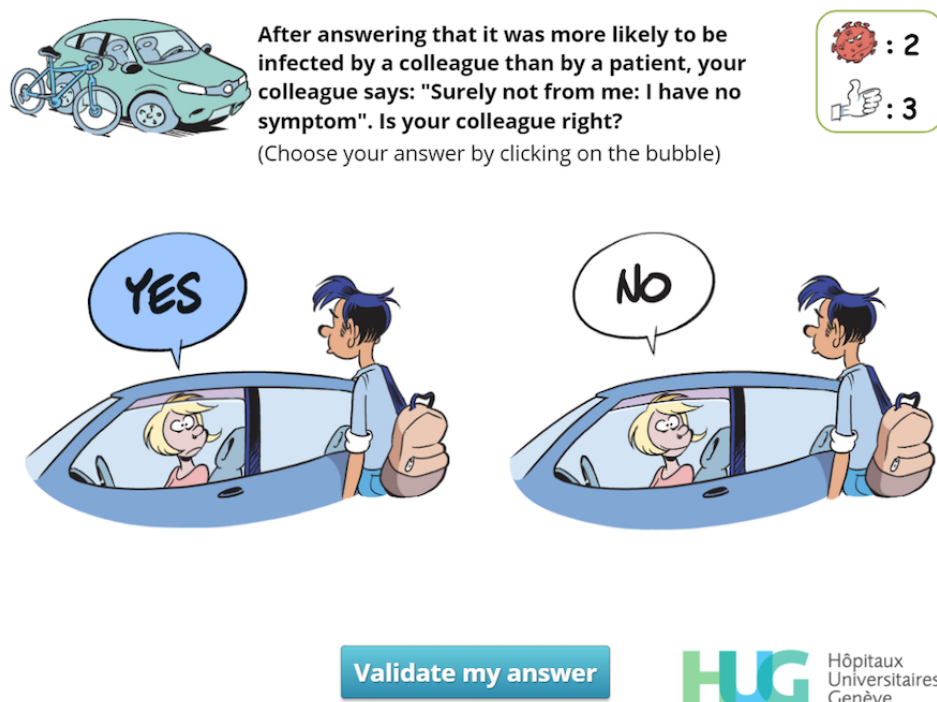
Figure 13. Multiple-choice interaction. The player has to identify the safety precautions that must be taken at work, including at the cafeteria.

You go to the cafeteria to share breakfast with your colleagues. What precautions should you take?

(Click on the correct answers. There can be more than one correct answer)

🦠 : 3
👍 : 3

Maintain an appropriate physical distance	**Physical distancing is unnecessary**	**Wear a face mask**
Wear gloves	**Wear an N95 mask**	**Practice hand hygiene on a regular basis**

Validate my answer

HUG Hôpitaux Universitaires Genève

Figure 14. Feedback. The player has failed to choose the expected answers at the first try and can click on "Retry" (retry) to have a second (and last) chance. The number of wrong answers is provided along with the number of correct answers missing from the original attempt.

You go to the cafeteria to share breakfast with your colleagues. What precautions should you take?

🦠 : 3
👍 : 3

Almost

You have selected **2 wrong answer(s)** and **1 correct answer(s) are missing**. Please try again!

Retry

Validate my answer

HUG Hôpitaux Universitaires Genève

Figure 15. Simple-choice interaction. The player must drag the "thumbs-up" on the image of the only colleague correctly wearing a face mask.
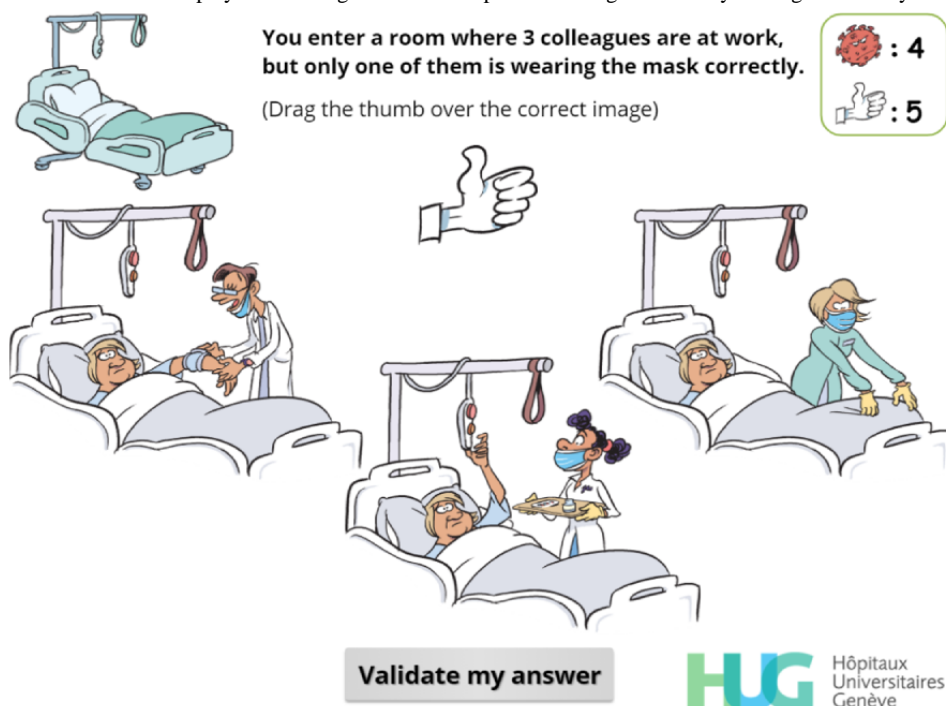
Figure 16. Multiple-choice interaction. The player must choose the situations in which eye protection should be worn.
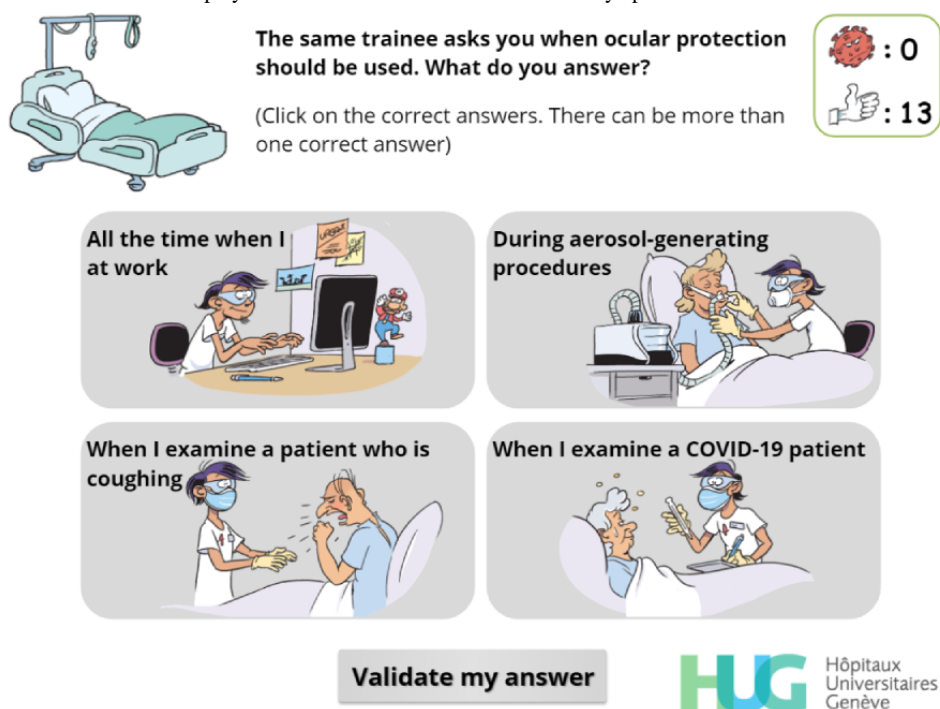
Figure 17. Multiple-choice interaction. The player must choose in which situations they should wear or change gloves. Illustrations change as the cursor hovers over the buttons on the left.
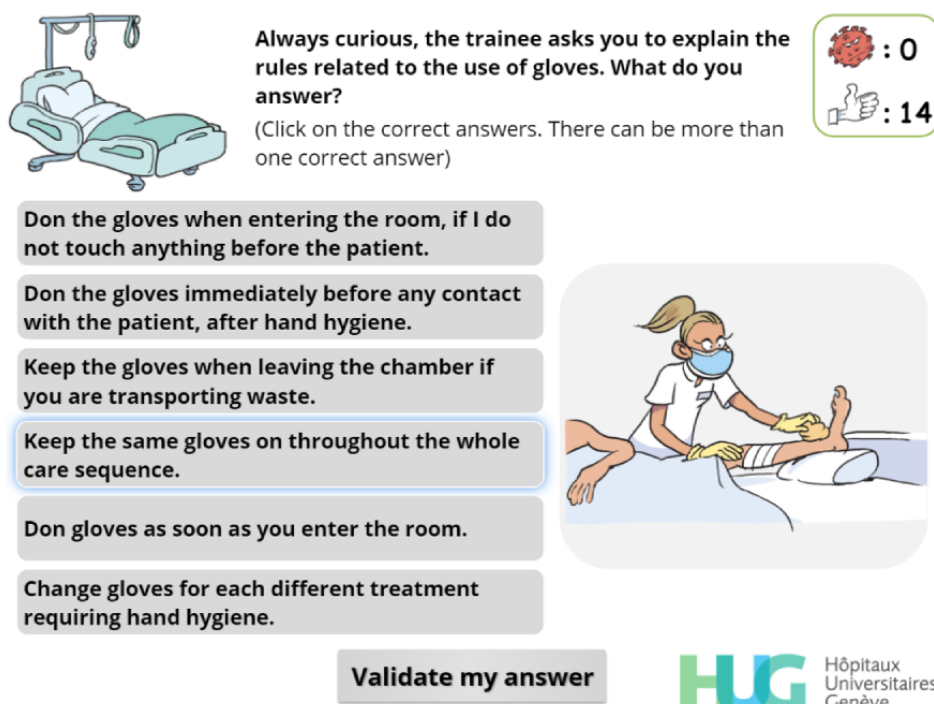
To keep track of the player's choices and to strengthen IPC messages, a virus counter was created. Whenever the player selects an answer that can lead to a possible contamination, a red virus image appears, and the virus count rises incrementally. Conversely, every time the player selects an answer that matches a desirable behavior, a positive token in the form of a "thumbs-up" image is awarded. If the player reaches a total of 5 viruses, a "game over" screen (Figure 18) appears. The user can then choose to decrease the virus count by exchanging "thumbs-up" tokens against viruses or to restart the level. We decided that the players should only restart the level, and not the entire game, in order to provide a more positive feedback and to avoid generating frustration. Moreover, a positive message is displayed at the end of each of the 4 levels. These messages are designed to strengthen the player's motivation to comply with IPC guidelines and are not influenced by the virus count.

Feedback mechanisms were used extensively and systematically in this serious game [32]. After a simple-choice interaction, the expected answer is immediately provided to the player along with an explanatory message (Figures 8 and 9). When multiple choice interactions are involved, players are allowed to adapt their answer if they were initially wrong. To help them find the correct answers, the number of wrong answers is provided along with the number of correct answers missing from their initial choice (Figure 14).

Figure 18. Game over screen. The player can choose to exchange their thumbs-up for viruses (1:1) or to restart the level and decrease the virus count by 2.
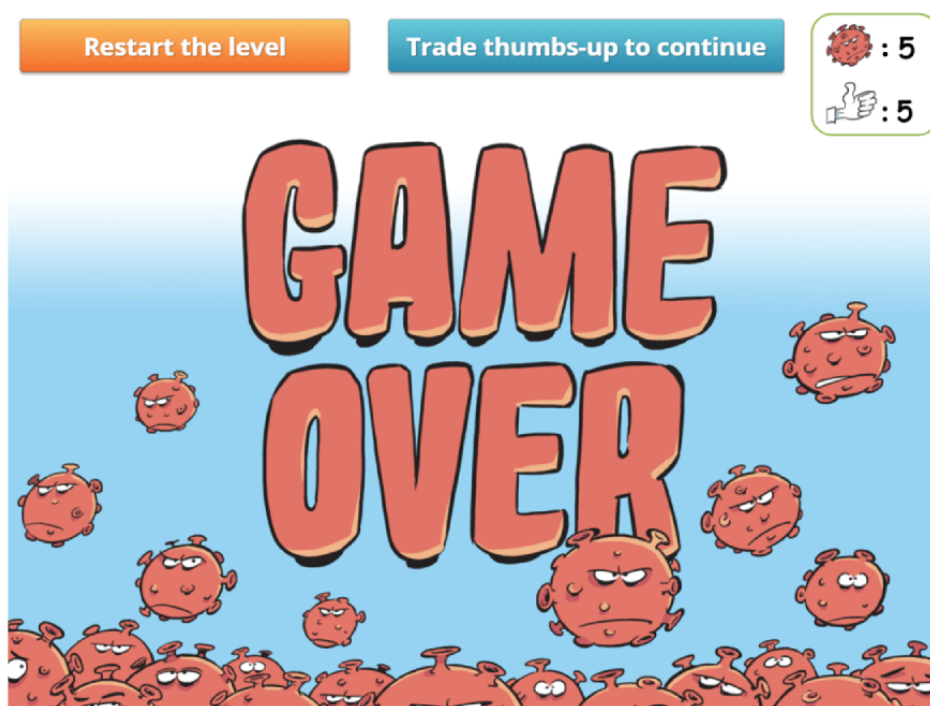
Game Development and Iterations

We first created a storyboard using PowerPoint 16 (Microsoft Corp). This storyboard, split into 4 game levels, was refined through multiple iterations until we deemed it advanced enough to begin the actual game development. A common and simple terminology was used to suit the broad target audience. We chose to develop the actual game under Articulate Storyline 3 (Articulate Global). This software is easy to use, embeds many elements that are essential for creating interactive content and serious games, and allows developers to export their creations to many different platforms using HTML5.

The first version of the game was developed using stick figures created by PresenterMedia (Eclipse Digital Imaging Inc). Figure 1 shows the welcome screen of this first version, and Figure 3 gives an example of a simple-choice interaction.

Although developers and IPC specialists found this first version to be both adequate and usable after having been refined through multiple iterations, they felt that it resembled a gamified e-learning module rather than a serious game. We therefore hired a well-known Swiss graphic designer to help us enhance the graphical aspects of the game and create a unique atmosphere. First, sketches in line with the multiple iterations concept of the SERES framework were produced by the graphic designer. These sketches were revised, adapted, and finally validated. They were then transformed into actual images that were embedded in the serious game. Figure 2 shows the updated (and final) version of the welcome screen, and Figure 4 displays the updated version a simple-choice interaction.

Game Validation

Each iteration of the game was tested by at least 3 authors whose feedback was systematically assessed and, when relevant and feasible, integrated into the next iteration of the game. The final version was validated by all the authors and by IPC specialists, some of whom act as consultants for the World Health Organization. Apart from the authors, all of whom are part of the medical or nursing staff, the game was tested by nonclinical staff from the Geneva University Hospitals training center before its release. These testing sessions showed that the game takes 15 minutes to complete on average.

The heuristic evaluation procedure described by Davids et al [49] was used to screen for usability issues. This led to the addition of blinking images in the welcome screen to facilitate identification of interactive elements.

Game Availability

This serious game can be freely accessed on the internet [50]. The corresponding author can be contacted to obtain a SCORM (shareable content object reference model) package of the game. We decided to make this package available to facilitate results and completion tracking on most current learning management system. This will allow for the tracking of completion and, on recent platforms, to gather detailed results down to the level of individual questions.

Discussion

Principal Findings

The first 3 stages of the SERES framework were successfully used to create the "Escape COVID-19" serious game. The content created meets the definition given by Gentry et al [20] in a recent systematic review (ie, a game created "for the serious purpose of providing health professions education via a digital device"). The boundaries separating serious games from other gamified educational tools are not easy to define [51]. Using the definition Payne et al [52] published in 2015, "Escape COVID-19" qualifies as a serious game as it primarily intends

to educate rather than entertain, contains a specific narrative, and includes the possibility of failure. In the context of the COVID-19 pandemic, new HCWs are being quickly recruited to face increasing health care demand [53]. Providing them regular face to face training in IPC is particularly challenging in terms of resources. Distance learning through serious games and other e-learning modalities offers an innovative and cost-effective alternative that can reach a large target population while avoiding the logistical challenges associated with face to face training.

Since HCWs perceive a higher likelihood of contracting COVID-19 than the general population [54] and as fatality rates have been shown to be substantial in this population [55,56], the impact of this serious game could be significant. Distance learning mechanisms have indeed been shown to be of particular value in the context of the current pandemic [57,58]. Still, there are many different e-learning modalities [59,60], and selecting the adequate modality to achieve the intended goal can be challenging. Our gamified e-learning module, which was intended to teach the adequate use of PPE to prehospital personnel, was found to be both useful and satisfactory by paramedics. It however failed to yield significant improvements regarding knowledge acquisition in this population [61]. Conversely, this e-learning modality significantly improved knowledge acquisition in student paramedics who were actively working in an ambulance service at the time of the study [62]. In both populations, however, this gamified module failed to enhance the acquisition of complicated procedures such as PPE donning or doffing [61,62]. Nevertheless, as the goal of this serious game is to promote and strengthen safe behaviors rather than to teach new content to HCWs, it might still prove effective in improving the application of IPC guidelines.

Though we chose to use Nicholson's RECIPE mnemonic for meaningful gamification [30], we were unable to apply it to its fullest extent. Indeed, the play element was only used in connection with one out of 10 acquisition objectives. Given the goal and design of this serious game, letting the player freely roam in a virtual environment was not an option. Moreover, we elected to force the player to follow a predetermined path, starting at home and ending in the ward. We made this choice to comply with the "flow" component of the engagement element [39]. When designing a serious game, using this component means that challenges should get harder as the player progresses. In the first level (at home), the challenges were relatively easy and the choice element was central. Conversely, in the last level, the player was asked to rebuild a complicated donning sequence and to make difficult choices regarding the kind of PPE that should be worn in specific situations. Therefore, forcing the player to follow a chronological order was relevant in this regard.

Limitations

The main limitation of this study is the current lack of evidence to support the use of this serious game. A longitudinal trial assessing the evolution of the rate of SARS-CoV-2 contamination in HCWs and in other hospital employees should be carried out to help assess the impact of this serious game. Assessing the added value of specific elements might however prove difficult, and multiple confounding factors (progressive dissemination of new guidelines, natural and unforeseeable evolution of the pandemic, etc) will make this more difficult.

Since knowledge regarding SARS-CoV-2 transmission mechanisms is evolving quickly, this serious game could be obsolete in a matter of months or less. This limitation is however relative as the platform used to develop this serious game is quite flexible and should allow us to rapidly update the game.

Finally, even though most screenshots are displayed in English, the original game was developed in French and the English translation has not been completed yet. It should however be completed by the end of 2020.

Future Work

While the creation of this serious game was considered successful, we cannot be certain of its actual impact on IPC practices and behavior. The SCORM package is available to any researcher or institution willing to deploy it on a learning management system, whether for research purposes or simply to make it available as an additional educational tool.

Following the release of version 2.08, the Geneva University Hospitals have deployed Escape COVID-19 and made it available to their 13,000 employees, including medical, nursing, and administrative and support staff. In the wake of this deployment, the Geneva Directorate of Health requested us to make this serious game available to nursing home employees. Indeed, long-term care facilities have been severely hit by the COVID-19 pandemic [46,63-65] and motivating their employees to enhance their IPC practices could be of significant benefit [66]. We naturally agreed and took this opportunity to initiate a web-based, triple-blinded randomized controlled trial [67]. The aim of this study is to determine whether Escape COVID-19 can enhance the intention to change IPC practices in nursing home employees.

Conclusion

The SERES framework was successfully used to create "Escape COVID-19," a serious game designed to promote safe behaviors among HCWs and other hospital employees during the COVID-19 pandemic. The impact of this game should now be assessed through a longitudinal trial.

Conflicts of Interest

None declared.

References

1. Coronavirus Resource Center. Johns Hopkins University & Medicine. 2020. URL: https://coronavirus.jhu.edu/ [accessed 2020-10-11]
2. Iacobucci G. Covid-19: GPs need extra support to withstand second wave, BMA warns. BMJ 2020 Oct 01;371:m3806. [doi: 10.1136/bmj.m3806] [Medline: 33004345]
3. Kaxiras E, Neofotistos G. Multiple Epidemic Wave Model of the COVID-19 Pandemic: Modeling Study. J Med Internet Res 2020 Jul 30;22(7):e20912 [FREE Full text] [doi: 10.2196/20912] [Medline: 32692690]
4. Xu S, Li Y. Beware of the second wave of COVID-19. The Lancet 2020 Apr;395(10233):1321-1322. [doi: 10.1016/s0140-6736(20)30845-x]
5. Xiao J, Fang M, Chen Q, He B. SARS, MERS and COVID-19 among healthcare workers: A narrative review. J Infect Public Health 2020 Jun;13(6):843-848 [FREE Full text] [doi: 10.1016/j.jiph.2020.05.019] [Medline: 32493671]
6. Zheng C, Hafezi-Bakhtiari N, Cooper V, Davidson H, Habibi M, Riley P, et al. Characteristics and transmission dynamics of COVID-19 in healthcare workers at a London teaching hospital. J Hosp Infect 2020 Oct;106(2):325-329 [FREE Full text] [doi: 10.1016/j.jhin.2020.07.025] [Medline: 32730771]
7. Chou R, Dana T, Buckley DI, Selph S, Fu R, Totten AM. Epidemiology of and Risk Factors for Coronavirus Infection in Health Care Workers. Annals of Internal Medicine 2020 Jul 21;173(2):120-136. [doi: 10.7326/m20-1632]
8. Stock AD, Bader ER, Cezayirli P, Inocencio J, Chalmers SA, Yassari R, et al. COVID-19 Infection Among Healthcare Workers: Serological Findings Supporting Routine Testing. Front Med (Lausanne) 2020 Aug 21;7:471. [doi: 10.3389/fmed.2020.00471] [Medline: 32974370]
9. Zhang Z, Liu S, Xiang M, Li S, Zhao D, Huang C, et al. Protecting healthcare personnel from 2019-nCoV infection risks: lessons and suggestions. Front Med 2020 Apr;14(2):229-231 [FREE Full text] [doi: 10.1007/s11684-020-0765-x] [Medline: 32212058]
10. Casanova L, Alfano-Sobsey E, Rutala WA, Weber DJ, Sobsey M. Virus transfer from personal protective equipment to healthcare employees' skin and clothing. Emerg Infect Dis 2008 Aug;14(8):1291-1293 [FREE Full text] [doi: 10.3201/eid1408.080085] [Medline: 18680659]
11. Barycka K, Torlinski T, Filipiak KJ, Jaguszewski M, Nadolny K, Szarpak L. Risk of self-contamination among healthcare workers in the COVID-19 pandemic. Am J Emerg Med 2020 Sep 28 [FREE Full text] [doi: 10.1016/j.ajem.2020.09.055] [Medline: 33008700]
12. Freeman D, Waite F, Rosebrock L, Petit A, Causier C, East A, et al. Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. Psychol. Med 2020 May 21:1-13. [doi: 10.1017/s0033291720001890]
13. Ahmed W, López Seguí F, Vidal-AQaball J, Katz MS. COVID-19 and the "Film Your Hospital" Conspiracy Theory: Social Network Analysis of Twitter Data. J Med Internet Res 2020 Oct 05;22(10):e22374 [FREE Full text] [doi: 10.2196/22374] [Medline: 32936771]
14. Bhagavathula AS, Aldhaleei WA, Rahmani J, Mahabadi MA, Bandari DK. Knowledge and Perceptions of COVID-19 Among Health Care Workers: Cross-Sectional Study. JMIR Public Health Surveill 2020 Apr 30;6(2):e19160 [FREE Full text] [doi: 10.2196/19160] [Medline: 32320381]
15. Gesser-Edelsburg A, Cohen R, Shahbari NAE, Hijazi R. A mixed-methods sequential explanatory design comparison between COVID-19 infection control guidelines' applicability and their protective value as perceived by Israeli healthcare workers, and healthcare executives' response. Antimicrob Resist Infect Control 2020 Sep 04;9(1):148 [FREE Full text] [doi: 10.1186/s13756-020-00812-8] [Medline: 32887658]
16. Maqbool A, Khan NZ. Analyzing barriers for implementation of public health and social measures to prevent the transmission of COVID-19 disease using DEMATEL method. Diabetes Metab Syndr 2020 Sep;14(5):887-892 [FREE Full text] [doi: 10.1016/j.dsx.2020.06.024] [Medline: 32563940]
17. Deci E, Ryan R. Handbook of Self-Determination Research. Rochester, NY: University of Rochester Press; 2002.
18. Tian W, Jiang W, Yao J, Nicholson CJ, Li RH, Sigurslid HH, et al. Predictors of mortality in hospitalized COVID-19 patients: A systematic review and meta-analysis. J Med Virol 2020 May 22;92(10):1875-1883 [FREE Full text] [doi: 10.1002/jmv.26050] [Medline: 32441789]
19. Bartsch SM, Ferguson MC, McKinnell JA, O'Shea KJ, Wedlock PT, Siegmund SS, et al. The Potential Health Care Costs And Resource Use Associated With COVID-19 In The United States. Health Aff (Millwood) 2020 Jun 23;39(6):927-935 [FREE Full text] [doi: 10.1377/hlthaff.2020.00426] [Medline: 32324428]
20. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious Gaming and Gamification Education in Health Professions: Systematic Review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]

21. Verschueren S, Buffel C, Vander Stichele G. Developing Theory-Driven, Evidence-Based Serious Games for Health: Framework Based on Research Community Insights. JMIR Serious Games 2019 May 02;7(2):e11565. [doi: 10.2196/11565] [Medline: 31045496]

22. Verschueren S, van Aalst J, Bangels A, Toelen J, Allegaert K, Buffel C, et al. Development of CliniPup, a Serious Game Aimed at Reducing Perioperative Anxiety and Pain in Children: Mixed Methods Study. JMIR Serious Games 2019 Jun 01;7(2):e12429 [FREE Full text] [doi: 10.2196/12429] [Medline: 31199333]

23. Suppan M, Gartner B, Golay E, Stuby L, White M, Cottet P, et al. Teaching Adequate Prehospital Use of Personal Protective Equipment During the COVID-19 Pandemic: Development of a Gamified e-Learning Module. JMIR Serious Games 2020 Jun 12;8(2):e20173 [FREE Full text] [doi: 10.2196/20173] [Medline: 32516115]

24. McCaffrey T. BrainSwarming: Because Brainstorming Doesn't Work. Harvard Business Review. 2014. URL: https://hbr.org/2014/06/brainswarming-because-brainstorming-doesnt-work [accessed 2020-11-30]

25. Houghton C, Meskell P, Delaney H, Smalle M, Glenton C, Booth A, et al. Barriers and facilitators to healthcare workers' adherence with infection prevention and control (IPC) guidelines for respiratory infectious diseases: a rapid qualitative evidence synthesis. Cochrane Database Syst Rev 2020 Apr 21;4:CD013582. [doi: 10.1002/14651858.CD013582] [Medline: 32315451]

26. Godlee F. Protect our healthcare workers. BMJ 2020 Apr 02:m1324. [doi: 10.1136/bmj.m1324]

27. Rhee C, Baker MA, Klompas M. The COVID-19 infection control arms race. Infect Control Hosp Epidemiol 2020 Nov 08;41(11):1323-1325 [FREE Full text] [doi: 10.1017/ice.2020.211] [Medline: 32381140]

28. Coronavirus - Recommandations pour les professionnels de la santé. Hôpitaux Universitaires de Genève. 2020. URL: https://www.hug.ch/coronavirus/recommandations-pour-professionnels-sante [accessed 2020-10-11]

29. Nouveau Coronavirus: Informations pour les professionnels de la santé. Federal Office of Public Health. 2020. URL: https://www.bag.admin.ch/bag/fr/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/information-fuer-die-aerzteschaft.html [accessed 2020-10-11]

30. Nicholson S. A RECIPE for Meaningful Gamification. In: Reiners T, Wood L, editors. Gamification in Education and Business. Cham: Springer; 2014:1-20.

31. Arnab S, Lim T, Carvalho MB, Bellotti F, de Freitas S, Louchart S, et al. Mapping learning and game mechanics for serious games analysis. Br J Educ Technol 2014 Jan 05;46(2):391-411. [doi: 10.1111/bjet.12113]

32. Kopp V, Stark R, Fischer M. Fostering diagnostic knowledge through computer-supported, case-based worked examples: effects of erroneous examples and feedback. Med Educ 2008 Aug;42(8):823-829. [doi: 10.1111/j.1365-2923.2008.03122.x] [Medline: 18564096]

33. Latimier A, Riegert A, Peyre H, Ly ST, Casati R, Ramus F. Does pre-testing promote better retention than post-testing? NPJ Sci Learn 2019 Sep 24;4(1):15 [FREE Full text] [doi: 10.1038/s41539-019-0053-1] [Medline: 31583117]

34. Raina S, Bernard L, Taylor B, Kaza S. Using eye-tracking to investigate content skipping: A study on learning modules in cybersecurity. 2016 Presented at: 2016 IEEE Conference on Intelligence and Security Informatics (ISI); Sept 28-30, 2016; Tucson, AZ, USA. [doi: 10.1109/ISI.2016.7745486]

35. Fouh E, Breakiron DA, Hamouda S, Farghally MF, Shaffer CA. Exploring students learning behavior with an interactive etextbook in computer science courses. Computers in Human Behavior 2014 Dec;41:478-485. [doi: 10.1016/j.chb.2014.09.061]

36. Koka A, Suppan L, Cottet P, Carrera E, Stuby L, Suppan M. Teaching the National Institutes of Health Stroke Scale to Paramedics (E-Learning vs Video): Randomized Controlled Trial. J Med Internet Res 2020 Jun 09;22(6):e18358 [FREE Full text] [doi: 10.2196/18358] [Medline: 32299792]

37. Drummond D, Hadchouel A, Tesnière A. Serious games for health: three steps forwards. Adv Simul (Lond) 2017 Feb 4;2(1):3 [FREE Full text] [doi: 10.1186/s41077-017-0036-3] [Medline: 29450004]

38. Snodgrass JG, Lacy MG, Francois Dengah HJ, Fagan J, Most DE. Magical flight and monstrous stress: technologies of absorption and mental wellness in Azeroth. Cult Med Psychiatry 2011 Mar 17;35(1):26-62. [doi: 10.1007/s11013-010-9197-4] [Medline: 21165683]

39. Csikszentmihalyi M. Finding Flow: The Psychology of Engagement With Everyday Life. New York: Basic Books; 1997.

40. Kohn A. Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes. Boston: Houghton Mifflin Harcourt; 1993.

41. Yin G, Jin H. Comparison of Transmissibility of Coronavirus Between Symptomatic and Asymptomatic Patients: Reanalysis of the Ningbo COVID-19 Data. JMIR Public Health Surveill 2020 May 28;6(2):e19464 [FREE Full text] [doi: 10.2196/19464] [Medline: 32442131]

42. Sommerstein R, Fux CA, Vuichard-Gysin D, Abbas M, Marschall J, Balmelli C, Swissnoso. Risk of SARS-CoV-2 transmission by aerosols, the rational use of masks, and protection of healthcare workers from COVID-19. Antimicrob Resist Infect Control 2020 Jul 06;9(1):100 [FREE Full text] [doi: 10.1186/s13756-020-00763-0] [Medline: 32631450]

43. Sharma S, Mishra M, Mudgal S. Efficacy of cloth face mask in prevention of novel coronavirus infection transmission: A systematic review and meta-analysis. J Educ Health Promot 2020;9(1):192 [FREE Full text] [doi: 10.4103/jehp.jehp_533_20] [Medline: 33015206]

44. Noorimotlagh Z, Mirzaee SA, Jaafarzadeh N, Maleki M, Kalvandi G, Karami C. A systematic review of emerging human coronavirus (SARS-CoV-2) outbreak: focus on disinfection methods, environmental survival, and control and prevention strategies. Environ Sci Pollut Res Int 2020 Oct 02 [FREE Full text] [doi: 10.1007/s11356-020-11060-z] [Medline: 33009614]

45. Li X, Sridhar S, Chan JFW. The Coronavirus Disease 2019 pandemic: how does it spread and how do we stop it? Curr Opin HIV AIDS 2020 Nov;15(6):328-335. [doi: 10.1097/COH.0000000000000650] [Medline: 32969975]

46. Roxby AC, Greninger AL, Hatfield KM, Lynch JB, Dellit TH, James A, et al. Outbreak Investigation of COVID-19 Among Residents and Staff of an Independent and Assisted Living Community for Older Adults in Seattle, Washington. JAMA Intern Med 2020 Aug 01;180(8):1101-1105 [FREE Full text] [doi: 10.1001/jamainternmed.2020.2233] [Medline: 32437547]

47. Caserman P, Hoffmann K, Müller P, Schaub M, Straßburg K, Wiemeyer J, et al. Quality Criteria for Serious Games: Serious Part, Game Part, and Balance. JMIR Serious Games 2020 Jul 24;8(3):e19037 [FREE Full text] [doi: 10.2196/19037] [Medline: 32706669]

48. Ravyse WS, Seugnet Blignaut A, Leendertz V, Woolner A. Success factors for serious games to enhance learning: a systematic review. Virtual Reality 2016 Sep 20;21(1):31-58. [doi: 10.1007/s10055-016-0298-4]

49. Davids M, Chikte U, Halperin M. An efficient approach to improve the usability of e-learning resources: the role of heuristic evaluation. Adv Physiol Educ 2013 Sep;37(3):242-248 [FREE Full text] [doi: 10.1152/advan.00043.2013] [Medline: 24022770]

50. Escape COVID-19 Serious Game. 2020. URL: https://coronavirus.anesth.ch/serious-game [accessed 2020-11-30]

51. Tan JW, Zary N. Diagnostic Markers of User Experience, Play, and Learning for Digital Serious Games: A Conceptual Framework Study. JMIR Serious Games 2019 Jul 16;7(3):e14620 [FREE Full text] [doi: 10.2196/14620] [Medline: 31313660]

52. Payne HE, Moxley VB, MacDonald E. Health Behavior Theory in Physical Activity Game Apps: A Content Analysis. JMIR Serious Games 2015 Jul 13;3(2):e4 [FREE Full text] [doi: 10.2196/games.4187] [Medline: 26168926]

53. Fagan JJ, Cairncross L, Biccard B, Fieggen G, Maswime S. COVID-19 exposes health worker shortages in the USA and UK, but nationalism and self-interest must not exploit medical workforces from low- and middle-income countries. S Afr Med J 2020 Apr 08;110(5):335. [doi: 10.7196/samj.2020.v110i5.14774] [Medline: 32657705]

54. Ali K, Whitebridge S, Jamal M, Alsafy M, Atkin S. Perceptions, Knowledge, and Behaviors Related to COVID-19 Among Social Media Users: Cross-Sectional Study. J Med Internet Res 2020 Sep 08;22(9):e19913. [doi: 10.2196/19913] [Medline: 32841153]

55. Lapolla P, Mingoli A, Lee R. Deaths from COVID-19 in healthcare workers in Italy-What can we learn? Infect Control Hosp Epidemiol 2020 May 15:1-2 [FREE Full text] [doi: 10.1017/ice.2020.241] [Medline: 32408922]

56. Hassanian-Moghaddam H, Zamani N, Kolahi A. COVID-19 pandemic, healthcare providers' contamination and death: an international view. Crit Care 2020 May 08;24(1):208 [FREE Full text] [doi: 10.1186/s13054-020-02938-y] [Medline: 32384910]

57. Chen CH, Mullen AJ. COVID-19 Can Catalyze the Modernization of Medical Education. JMIR Med Educ 2020 Jun 12;6(1):e19725 [FREE Full text] [doi: 10.2196/19725] [Medline: 32501809]

58. Remtulla R. The Present and Future Applications of Technology in Adapting Medical Education Amidst the COVID-19 Pandemic. JMIR Med Educ 2020 Jul 17;6(2):e20190. [doi: 10.2196/20190] [Medline: 32634107]

59. Wheeler S. e-Learning and Digital Learning. In: Seel NM, editor. Encyclopedia of the Sciences of Learning. Boston, MA: Springer; 2012.

60. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional Design Variations in Internet-Based Learning for Health Professions Education: A Systematic Review and Meta-Analysis. Academic Medicine 2010 May;85(5):909-922. [doi: 10.1097/ACM.0b013e3181d6c319] [Medline: 20520049]

61. Suppan L, Abbas M, Stuby L, Cottet P, Larribau R, Golay E, et al. Effect of an E-Learning Module on Personal Protective Equipment Proficiency Among Prehospital Personnel: Web-Based Randomized Controlled Trial. J Med Internet Res 2020 Aug 21;22(8):e21265 [FREE Full text] [doi: 10.2196/21265] [Medline: 32747329]

62. Suppan L, Stuby L, Gartner B, Larribau R, Iten A, Abbas M, et al. Impact of an e-learning module on personal protective equipment knowledge in student paramedics: a randomized controlled trial. Antimicrob Resist Infect Control 2020 Nov 10;9(1):185. [doi: 10.1186/s13756-020-00849-9] [Medline: 33168097]

63. Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. N Engl J Med 2020 May 28;382(22):2081-2090. [doi: 10.1056/nejmoa2008457] [Medline: 32329971]

64. McMichael TM, Currie DW, Clark S, Pogosjans S, Kay M, Schwartz NG, Public Health–SeattleKing County, EvergreenHealth,CDC COVID-19 Investigation Team. Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington. N Engl J Med 2020 May 21;382(21):2005-2011 [FREE Full text] [doi: 10.1056/NEJMoa2005412] [Medline: 32220208]

65. ECDC Public Health Emergency Team, Danis K, Fonteneau L, Georges S, Daniau C, Bernard-Stoecklin S, et al. High impact of COVID-19 in long-term care facilities, suggestion for monitoring in the EU/EEA, May 2020. Euro Surveill 2020 Jun;25(22) [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.22.2000956] [Medline: 32524949]

66. Szczerbińska K. Could we have done better with COVID-19 in nursing homes? Eur Geriatr Med 2020 Aug 16;11(4):639-643 [FREE Full text] [doi: 10.1007/s41999-020-00362-7] [Medline: 32676866]

67. Suppan L, Abbas M, Catho G, Stuby L, Regard S, Harbarth S, et al. Impact of a Serious Game on the Intention to Change Infection Prevention and Control Practices in Nursing Homes during the COVID-19 Pandemic: Protocol for a Web-Based, Randomized Controlled Trial. JMIR Preprints. Preprint posted online November 2020 2020 [FREE Full text] [doi: 10.2196/preprints.25595]

Abbreviations

e-learning: electronic learning
HCW: health care worker
IPC: infection prevention and control
MeSH: Medical Subject Headings
PPE: personal protective equipment
RECIPE: reflection, engagement, choice, information, play, exposition
SCORM: shareable content object reference model

XSL•FO
RenderX

<u>Original Paper</u>

Physicians' Perceptions of a Situation Awareness–Oriented Visualization Technology for Viscoelastic Blood Coagulation Management (Visual Clot): Mixed Methods Study

Tadzio Raoul Roche[1*]; Sadiq Said[1*]; Julian Rössler[1], MD; Malgorzata Gozdzik[1]; Patrick Meybohm[2], MD; Kai Zacharowski[3], MD; Donat R Spahn[1], MD; Christoph B Nöthiger[1], MD; David W Tscholl[1], MD

[1]Institute of Anesthesiology, University of Zurich and University Hospital Zurich, Zürich, Switzerland

[2]Department of Anesthesiology, University Hospital Würzburg, Würzburg, Germany

[3]Department of Anesthesiology, Intensive Care Medicine and Pain Therapy, University Hospital Frankfurt, Frankfurt am Main, Germany

[*]these authors contributed equally

Corresponding Author:
David W Tscholl, MD
Institute of Anesthesiology
University of Zurich and University Hospital Zurich
Raemistrasse 100
Zürich, 8091
Switzerland
Phone: 41 44 253 94 51
Fax: 41 44 255 44 09
Email: david.tscholl@usz.ch

Abstract

Background: Viscoelastic tests enable a time-efficient analysis of coagulation properties. An important limitation of viscoelastic tests is the complicated presentation of their results in the form of abstract graphs with a multitude of numbers. We developed Visual Clot to simplify the interpretation of presented clotting information. This visualization technology applies user-centered design principles to create an animated model of a blood clot during the hemostatic cascade. In a previous simulation study, we found Visual Clot to double diagnostic accuracy, reduce time to decision making and perceived workload, and improve care providers' confidence.

Objective: This study aimed to investigate the opinions of physicians on Visual Clot technology. It further aimed to assess its strengths, limitations, and clinical applicability as a support tool for coagulation management.

Methods: This was a researcher-initiated, international, double-center, mixed qualitative-quantitative study that included the anesthesiologists and intensive care physicians who participated in the previous Visual Clot study. After the participants solved six coagulation scenarios using Visual Clot, we questioned them about the perceived pros and cons of this new tool. Employing qualitative research methods, we identified recurring answer patterns, and derived major topics and subthemes through inductive coding. Based on them, we defined six statements. The study participants later rated their agreement to these statements on five-point Likert scales in an online survey, which represented the quantitative part of this study.

Results: A total of 60 physicians participated in the primary Visual Clot study. Among these, 36 gave an interview and 42 completed the online survey. In total, eight different major topics were derived from the interview field note responses. The three most common topics were "positive design features" (29/36, 81%), "facilitates decision making" (17/36, 47%), and "quantification not made" (17/36, 47%). In the online survey, 93% (39/42) agreed to the statement that Visual Clot is intuitive and easy to learn. Moreover, 90% (38/42) of the participants agreed that they would like the standard result and Visual Clot displayed on the screen side by side. Furthermore, 86% (36/42) indicated that Visual Clot allows them to deal with complex coagulation situations more quickly.

Conclusions: A group of anesthesia and intensive care physicians from two university hospitals in central Europe considered Visual Clot technology to be intuitive, easy to learn, and useful for decision making in situations of active bleeding. From the responses of these possible future users, Visual Clot appears to constitute an efficient and well-accepted way to streamline the decision-making process in viscoelastic test–based coagulation management.

XSL•FO
RenderX

Introduction

For optimal perioperative bleeding management, a quick and reliable assessment of the patient's blood coagulation function is of utmost importance [1-3]. As laboratory coagulation testing can take more than an hour [4] to produce results, viscoelastic point-of-care devices, such as rotational thromboelastometry (ROTEM, Instrumentation Laboratory/Werfen) and thromboelastography (Haemonetics) play an important role in guiding hemostatic interventions in a time-efficient manner [5]. Previous studies showed that the utilization of viscoelastic testing can improve the patient outcome in those with bleeding trauma [3,6-9]. Hence, the European guidelines on trauma management recommend viscoelastic testing to enable goal-directed bleeding management [10]. Furthermore, viscoelastic point-of-care coagulation testing reduces the transfusion of allogeneic blood products in trauma [2,10]. Previous studies also showed the benefits of its use in cardiac [1,11,12], transplant [13], neuro [14], and pediatric surgery [5].

However, the abstract presentation of viscoelastic test results complicates its handling in clinical routine. This in turn shows the demand for a tool that facilitates the interpretation of viscoelastic readings. To serve this apparent need, Visual Clot technology [15] was developed. Visual Clot is an alternative situation awareness–oriented visualization technology created by adhering to the principles of user-centered design [16]. Instead of abstract shapes and a multitude of numbers, as in the conventional presentation of the ROTEM results, Visual Clot technology displays a three-dimensional animated model of a blood clot that corresponds to the real phenomena one would see when looking at the blood clot through an electron microscope. In an international dual-center study [15], Visual Clot technology enabled the participating physicians to make twice as many correct diagnoses quicker and with improved diagnostic confidence and reduced perceived cognitive workload compared with standard viscoelastic test results.

This study aimed to explore and capture the opinions of the 60 anesthesiologists and intensive care specialists from the previous Visual Clot study [15] on this new visualization technology. The results will help define the strengths and limitations of this visual decision aid. Further, this study examined the applicability of Visual Clot as an additional tool for coagulation management from the viewpoint of its potential future users.

Methods

Approval and Consent

The leading ethics committee (the Cantonal Ethics Committee of the Canton of Zurich in Switzerland) reviewed the study protocol and issued a declaration of no objection (Business Management System for Ethics Committees Number 2018-00933). We obtained written informed consent from each participant for the use of their data.

Study Design

We conducted this study at the University Hospital Frankfurt (UKF) in Germany and the University Hospital Zurich (USZ) in Switzerland. We included anesthesiologists and intensive care physicians from these two hospitals. Both study centers routinely use ROTEM-guided hemostatic resuscitation.

This study is a researcher-initiated, international, dual-center, mixed qualitative-quantitative study about the opinions of anesthesia and intensive care physicians regarding Visual Clot technology. The methods employed included interviewing physicians after their first contact with Visual Clot technology and jointly generating field notes based on their answers. Further, we created statements derived from these field note responses and asked the same group of physicians to rate the statements in an online survey.

Previous Visual Clot Study and Participant Interviews

We held interviews for the qualitative part of this study at the end of the data collection sessions of the previously published Visual Clot study. In the mentioned study, the participating physicians evaluated 12 coagulation scenarios in randomized order. They solved each scenario twice, once only observing Visual Clot and once only using the conventional viscoelastic test results. Before the testing took place, the participants received brief individual training and subsequently solved six scenarios using Visual Clot to become familiar with its features. We based the individual training on a Microsoft PowerPoint presentation (Microsoft Corporation). This aimed to explain all visualizations of Visual Clot and enabled the participants to ask questions in case of uncertainties. Multimedia Appendix 1 shows the training video for the current version of Visual Clot. Figure 1 explains and illustrates some visualizations of Visual Clot technology.

In an undisturbed environment, we asked the physicians about the advantages and disadvantages of Visual Clot. There was no further guidance by the investigators, who all had previous experience with qualitative research. While the physicians were encouraged to freely verbalize their thoughts, the data collector simultaneously took field notes in a Microsoft Word (Microsoft Corporation) document. After the interview session, the field notes were presented to the participants. They were asked whether they agreed with them and were encouraged to determine final adjustments and changes.

XSL•FO

RenderX

Figure 1. Graphics illustrating Visual Clot technology. A: The different coagulation components that Visual Clot represents as either present or absent (when dashed). B: A healthy Visual Clot with all coagulation elements present in sufficient quantity. C: A bleeding Visual Clot with heparin effect. D: A bleeding Visual Clot indicating the absence of plasmatic factors.
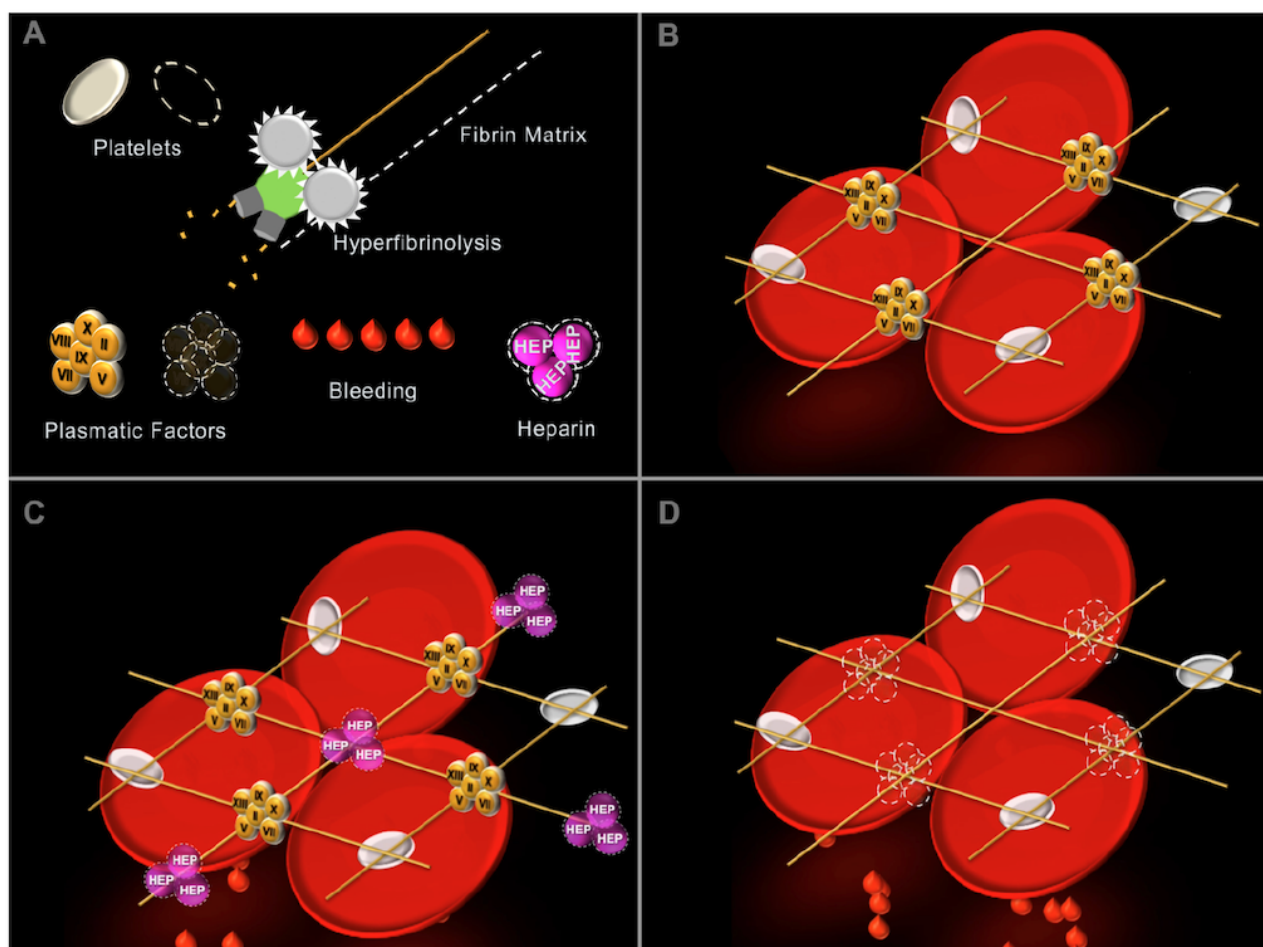
Qualitative Analysis

We used a template approach in the qualitative analysis [17]. This involves grouping topics using a coding template. The topics are often predefined in advance. During further data analysis, new topics are added or existing ones are revised [17]. First, we translated the original answers from German to English using DeepL [18]. All translated field notes are available in Multimedia Appendix 2. After translation, we used Microsoft Word to identify the most common words and created a tag cloud (Figure 2) as its quantitative graphic illustration using Wordle [19]. We ignored common English words (the, and, to, etc) and unified word groups with the same word root. The two physicians and study authors TRR and DWT were both involved in the qualitative analysis of the field notes. TRR is a resident anesthesiologist who did not partake in the interviewing process.

DWT is a senior anesthesiologist with previous experience in qualitative and patient safety research. Adhering to the criteria for reporting qualitative research [20], TRR and DWT developed a coding template as a rating system, using both word count and inductive coding based on recurring topics in the field note responses. We modified the initial coding template by successive reading, coding of the data, and discussions with other authors. After multiple data coding events, we agreed upon a final template. The two study authors mentioned above then independently rated all field notes using an illustrated final coding template, which has been provided in Figure 3. Interrater reliability was calculated to investigate the consistency of the implementation of the rating system. In cases of disagreement between these two examiners, a final code for the respective field note response was jointly determined.

Figure 2. A tag cloud as a quantitative graphic illustration of the most common words in the interview field notes. This word cloud was created using Wordle.net. Common English words (the, and, to, etc) are not displayed. Frequently occurring terms have larger font sizes.
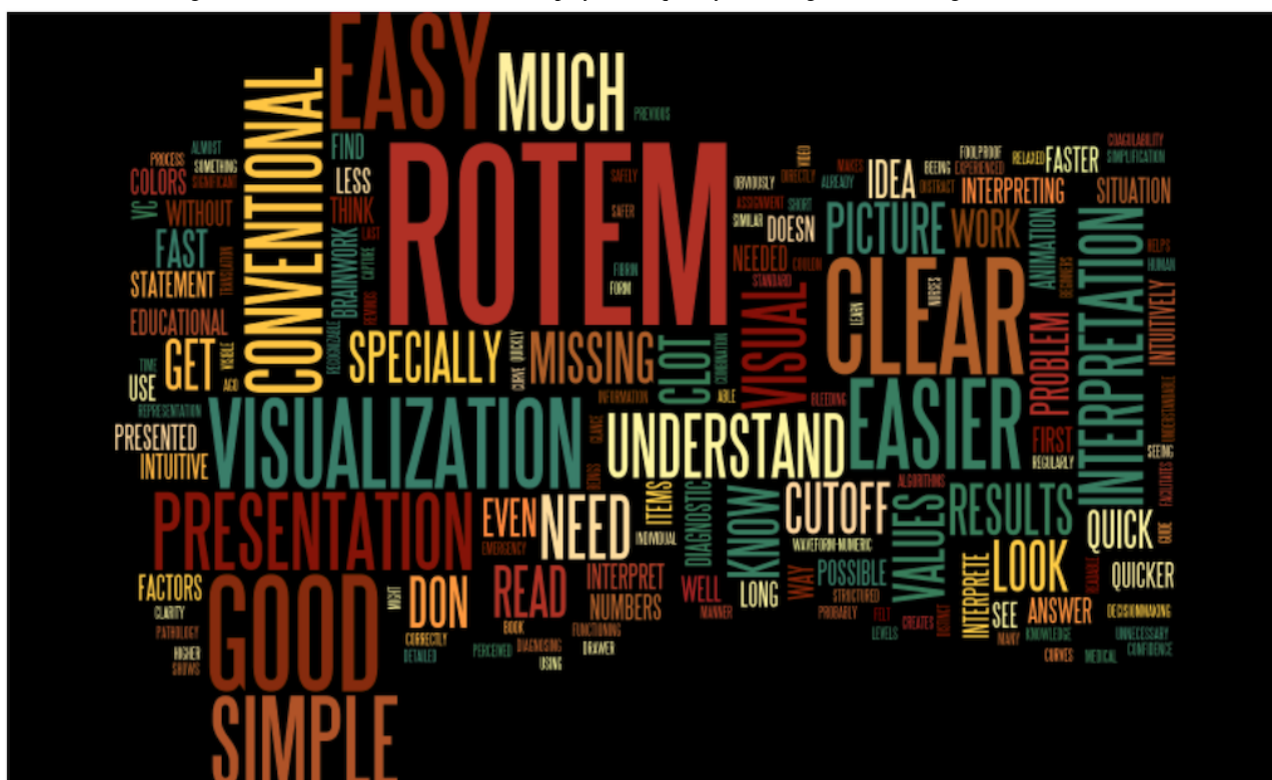
Figure 3. The coding template displaying the major topics and associated subthemes. We generated this through deductive coding via word count and inductive free coding of recurring topics in the interview field notes. A total of 36 participants were interviewed.
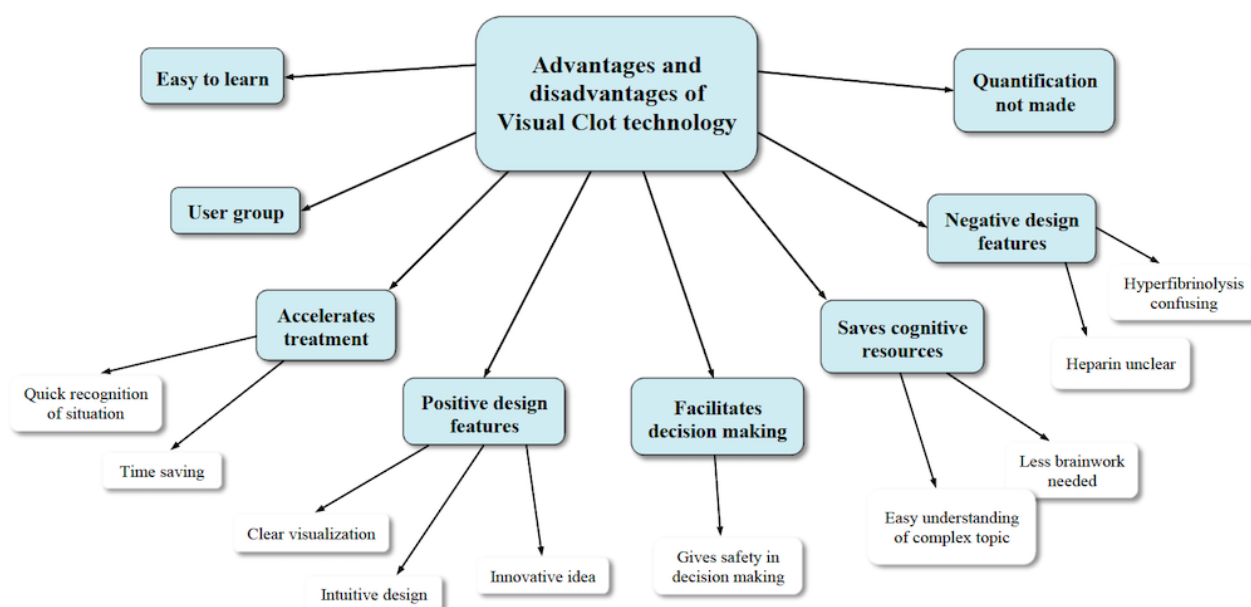
Quantitative Analysis

Literature provides several arguments why qualitative data could be combined with a quantitative analysis. One of the reasons for this is that quantitative data can help to generalize and confirm specific observations made in a qualitative analysis [21]. To further examine the qualitative research part of this study and Visual Clot applicability, we performed a quantitative assessment of the derived statements using a web-based online survey. We defined six partially generalized statements and

asked the participants to rate them on five-point Likert scales with responses ranging from "strongly disagree" to "strongly agree." Out of all six statements, four directly refer to the previously identified major topics. We specifically added two additional statements to the questioning that we considered necessary. These two aimed to obtain a deeper understanding of Visual Clot technology applicability. Using SurveyMonkey (SVMK Inc), we created the open online survey and tested its usability before sending the study link via email to all physicians who participated in the previous Visual Clot study and still worked at the respective institutions. We informed the participants that the survey takes about 1 minute to fill out. Participation was voluntary, no incentives were offered, and no personal identifying information was collected. The participants were able to review and change their answers before completion. A single reminder to complete the survey was sent after 2 weeks. Multimedia Appendix 3 displays the translated wording of the survey invitation announcement as well as the reminder mail. We completed the data collection 1 week after the reminder period expired.

Statistical Analysis

We provide the quantitative analysis data of the online survey as median and interquartile range for all statements. Using the Wilcoxon signed-rank test, we evaluated the difference between the median for each of the statements and the neutral answer. We considered $P<.05$ to indicate statistical significance. The qualitative part of this study aimed to identify the pros and cons using the new visualization technology. With the evaluated statements of the online survey, we aimed to further quantify

the agreement or disagreement. This procedure additionally increases the relevance of the statements made.

Data Sharing Statement

The translated interview field notes are available in Multimedia Appendix 2. We report all other data in this manuscript.

Results

Study and Participant Characteristics

This study included the same participants as in the previous Visual Clot study [15]. A total of 60 anesthesia and intensive care physicians participated. Half of them were working at the USZ and the other half at the UKF. After collecting data from the first 24 participants of the Visual Clot study, we decided to conduct systematic interviews as we found that the participants provided very informative feedback on the technology. We interviewed all of the participants at the UKF and six participants at the USZ (total of 36 out of 60 [60%]). Regarding the online survey, we present the results according to the checklist for reporting results of internet e-surveys [22]. We sent out the study link as an invitation to all physicians who still worked at the respective institutions. We checked the IP address displayed on the SurveyMonkey website to identify potential duplicate entries from the same user. Since the computers in the hospitals are shared and the workspace usually only changes daily, we considered a difference of at least 24 hours as appropriate for entries from the same IP address. Of all invitations sent out for the online survey, the participation rate was 86% (42/49). Table 1 further outlines the study and participant characteristics in detail.

Table 1. Study and participant characteristics.

Characteristics	Value
Study characteristics	
Total number of participants, n	60
Total number of interviewed participants (field notes) (N=60), n (%)	36 (60)
Online survey participation rate (N=49), n (%)	42 (86)
Online survey completion rate (N=42), n (%)	42 (100)
Participant characteristics	
Female sex (N=60), n (%)	23 (38)
Senior physicians (N=60), n (%)	35 (58)
Resident physicians (N=60), n (%)	25 (42)
Anesthesia experience in years, mean (IQR)	8 (4-11)
Number of annually interpreted ROTEM[a] readings per physician, mean (IQR)	40 (10-53)

[a]ROTEM: rotational thromboelastometry.

Part I: Qualitative Analysis of Interview Answers

The analysis of word count showed the following 12 most frequently occurring words or word groups in the participants' field note responses: visual/visualization/visually (24/36 participants, 67%), easy/easier/easily (18/36 participants, 50%), ROTEM (18/36 participants, 50%), interpretation/interpret/interpreting (13/36 participants, 36%),

presentation (12/36 participants, 33%), good (11/36 participants, 31%), clear (11/36 participants, 31%), simple/simplification/simplicity (10/36 participants, 28%), understand/understood/understandable (10/36 participants, 28%), quick/quicker/quickly (6/36 participants, 17%), cut-off/cut-off values (6/36 participants, 17%), and hyperfibrinolysis (6/36 participants, 17%). Figure 2 provides

XSL•FO
RenderX

the tag cloud created from the word count of the field note responses.

Through inductive free coding, we identified eight major topics with associated subthemes. Figure 3 displays the generated coding template, which was used as a coding system. The two study authors TRR and DWT rated all 36 field note responses

independently. A total of 131 codes were assessed. Interrater reliability was 71%, with a Cohen kappa of 0.665. After the first run, TRR and DWT discussed all differences and agreed on a coding in the case of disagreement. The second run showed an interrater agreement of 100% between the coding of TRR and DWT. Table 2 outlines all major topics with participant counts, percentages, and examples.

Table 2. The major topics with participant count, percentages, and examples (N=36)

Major topic	Examples
Positive design features (29 participants, 81%)	Participant #15: Intuitive kind of presentation.
	Participant #7: Clear presentation, no unnecessary information to distract.
	Participant #1: Obviously, something like this was missing.
Facilitates decision making (17 participants, 47%)	Participant #19: Creates clarity in emergency situations.
	Participant #30: Perceived diagnostic confidence is higher.
Saves cognitive resources (15 participants, 42%)	Participant #9: Don't have to think as much as with conventional ROTEM.
	Participant #16: Couldn't imagine, that it could be presented in such a simple way!
Easy to learn (14 participants, 39%)	Participant #05: Even without extensive previous education.
	Participant #33: Easy. Self-explaining.
	Participant #13: Fast learning.
Accelerates treatment (13 participants, 36%)	Participant #19: You can see immediately where the problem is.
	Participant #24: Quickly realize what the problem is.
	Participant #26: Faster with the Visual Clot.
	Participant #30: You get the answer much faster.
User groups (9 participants, 25%)	Participant #25: Can be used safely by all educational levels.
Quantification not made (17 participants, 47%)	Participant #14: No quantitative data or information.
Negative design features (17 participants, 47%)	Participant #22: Missing fibrin is not detected and understood easily.
	Participant #23: Heparin effect is a little tricky to interpret.

Themes

Positive Design Features

Of the 36 participants, 29 (81%) made comments that fit the major topic positive design features. After additional inductive free coding, this topic was further divided into the three subthemes "intuitive design," "clear visualization," and "innovative idea." Some participants perceived the visualization as intuitive, while others pointed out its clear visualization of standard ROTEM results. Participant #7 stated that there was no unnecessary information to distract. Others indicated Visual Clot as innovative, as they perceived this idea to be missing.

Facilitates Decision Making

Of the 36 participants, 17 (47%) noticed during the interview process how Visual Clot technology facilitated their decision making on fulfilling the given task. After free inductive coding, this major topic was broken down into the subtheme "safety in decision making" as some of these participants mentioned this aspect. Participant #2 stated how using Visual Clot made him less afraid of getting involved with the ROTEM analysis. Additionally, participant #19 pointed out that the simplicity of Visual Clot technology created clarity in emergency situations.

Saves Cognitive Resources

Of the 36 participants, 15 (42%) indicated that using Visual Clot, the user saves cognitive resources compared with standard ROTEM readings. The two subthemes "less brainwork needed" and "easy understanding of complex topic" are derived from this major topic. Some participants mentioned that they did not have to think as much using Visual Clot. Another participant reasoned that using Visual Clot, he did not need to memorize the cut-off values.

Easy to Learn

Of the 36 participants, 14 (39%) remarked that Visual Clot technology is easy to learn. Participant #25 felt that after a short instruction, one is already able to use Visual Clot correctly.

Accelerating Treatment

Of the 36 participants, 13 (36%) found Visual Clot to accelerate treatment. This major topic was subdivided into the aspects "quick recognition of situation" and "time saving." Some participants mentioned that using Visual clot, most of the information is visible at a single look. Further, it helps them to interpret the coagulation status faster.

User Groups

Of the 36 participants, 9 (25%) made comments that fit the major topic user groups. All of them shared the opinion that

Visual Clot would especially help an inexperienced user in the interpretation of ROTEM readings. Participant #25 noted that the visualization technology can be safely used at all educational levels.

Quantification Not Made

Of the 36 participants, 17 (47%) had concerns about the missing quantitative data using Visual Clot. Participant #16 suggested to show the visualization technology and the conventional ROTEM results next to each other to overcome this issue. Some mentioned how Visual Clot does not show the extent of the coagulation disorder. Participant #2 was concerned about whether this technology was reliable.

Negative Design Features

Of the 36 participants, 17 (47%) made comments that fit the major topic "negative design features." Problems in differentiating between hyperfibrinolysis and a low fibrinogen state were mentioned by several participants. Moreover, the heparin effect was difficult to interpret. Participant #1 found the heparin visualization confusing.

Part II: Quantitative Analysis of Statements Rated in the Online Survey

We illustrate the ratings of the six statements from the online survey in Figure 4. The sample medians of statements two to six all differed statistically significantly from neutral ($P<.001$). The first statement showed that half of the participants

considered the use of ROTEM readings in the management of coagulopathies difficult, whereas the other half did not.

The six statements with the evaluation of the online survey rating results are as follows:

(1) "I find the interpretation of the conventional presentation of the ROTEM and the management of coagulopathies difficult." This statement does not differ statistically significantly from neutral ($P=.54$).

(2) "I find Visual Clot intuitive and I could learn the interpretation easily." Of the 42 participants, 39 (93%) agreed or strongly agreed with this statement ($P<.001$).

(3) "Visual Clot allows me to quickly grasp complex clotting situations." Of the 42 participants, 36 (86%) agreed or strongly agreed with this statement ($P<.001$).

(4) "I think Visual Clot is helpful in the decision-making process for a targeted therapy." Of the 42 participants, 34 (81%) agreed or strongly agreed with this statement ($P<.001$).

(5) "I required more cognitive resources for the interpretation of Visual Clot than for the interpretation of the conventional ROTEM." Of the 42 participants, 1 (2%) agreed with this statement ($P<.001$).

(6) "I would find it helpful to have the visualization of Visual Clot and the conventional ROTEM on the screen next to each other." Of the 42 participants, 38 (90%) agreed or strongly agreed with this statement ($P<.001$).

Figure 4. Online survey results presented as donut parts of whole charts with the number of participants who chose a particular category (N=42). We present the results as medians and interquartile ranges and provide P values. ROTEM: rotational thromboelastometry.
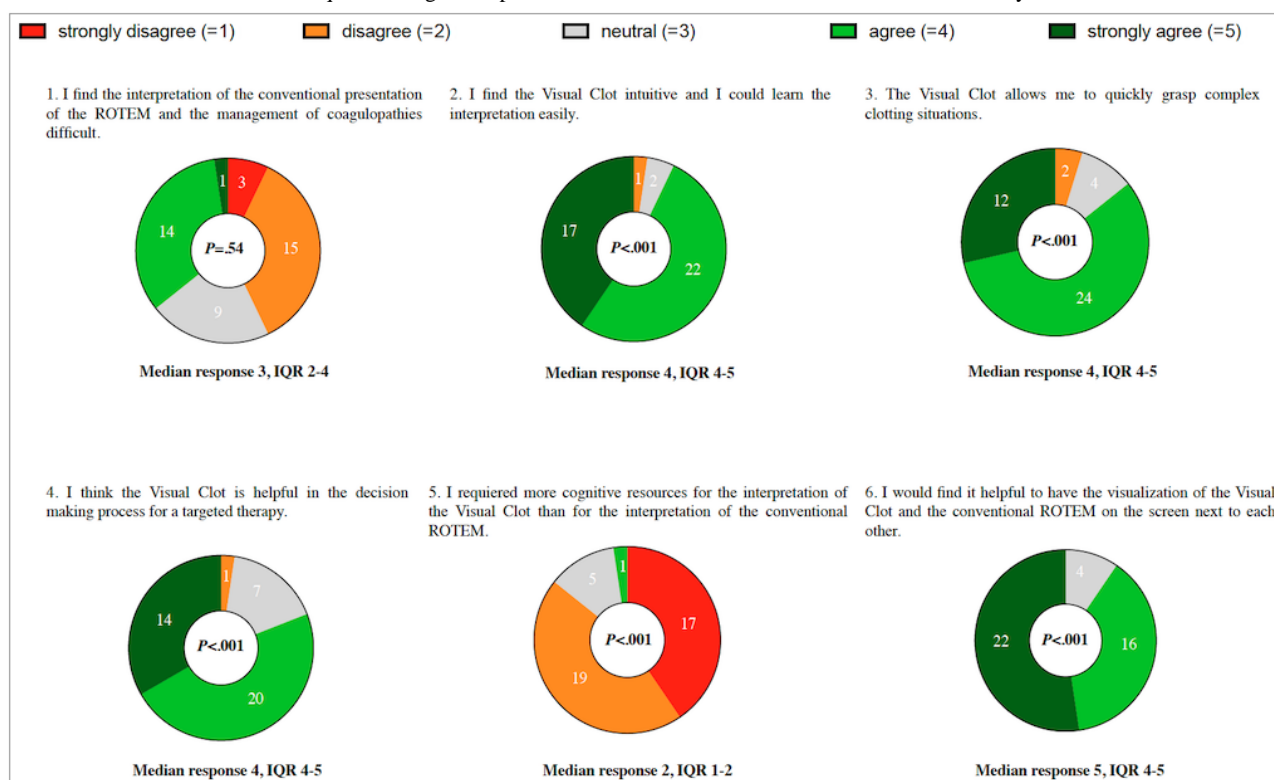
Discussion

Principal Findings

This mixed-methods study analyzed the opinions of physicians regarding Visual Clot, a new situation-awareness oriented visualization technology for viscoelastic coagulation management. The main findings were that the participants perceived Visual Clot as intuitive, easy to learn, and helpful in the decision-making process for ROTEM-guided coagulation management. Further, they found that Visual Clot gave them a good overview of the clotting situation. The main criticism concerned its missing quantification. The participants preferred this visualization on a screen next to the conventional ROTEM readings. In an environment where health care providers are confronted with increasingly polymorbid patients and complicated diagnostic tools [23,24], more efficient assistive technologies leading to safer transmission of critical information will be of lasting importance for human performance and patient safety.

Our study showed a high level of physician acceptance and satisfaction with the new tool. Eighty-one percent (34/42) of the participants regarded the animated blood clot as useful in the decision-making process for coagulation management. Only one participant (N=42, 2%) agreed to the statement "I required more cognitive resources for the interpretation of Visual Clot than for the interpretation of conventional ROTEM." The participant responses corresponded to the reactions to the similarly well-accepted Visual Patient technology [25]. In that assessment, only 11% (4/38) of the questioned subjects considered that technology not helpful in patient monitoring after their first contact with it [26]. Visual Patient is a situation awareness–oriented visualization for patient vital signs with comparable effect size in information transfer improvement as that of Visual Clot [27-30]. Ninety-three percent (39/42) of the participants considered Visual Clot intuitive to interpret and easy to learn. In the Visual Patient study [26], 82% (31/38) of the respondents attributed the same characteristics to Visual Patient. The higher complexity of Visual Patient with far more visualizations may have caused the 10% difference in this aspect between the two technologies. Intuition enables humans to apply a new tool utilizing mainly unconscious processing and previously learned experiential knowledge [31]. Cognitive ease when learning a new technology is crucial for user acceptance [32]. Indeed, previous medical visualization technologies failed because they were too difficult to understand [33,34]. We drafted Visual Clot and its visualizations based on principles of logic, human-computer interaction, and results of prior work in medical interface and user-centered design [16,34]. According to the theory by Wittgenstein [35], a coherent image or model has a meaningful commonality with the reality it is intended to reflect. To achieve this, we designed Visual Clot as a model of a blood clot. This philosophy is in line with results from the study by Wachter et al, which show that an anatomically correct interface is particularly intuitive [33]. A previous National Aeronautics and Space Administration publication outlines the various hierarchical levels of information representation [36]. It highlights the "order of wholeness" achieved by integrating the required information into a single display as the highest

level of such presentation. In order to increase the cognitive receptiveness of information, we intentionally animated Visual Clot in a playful way [31,32]. This enhances user accessibility to the otherwise dry data-driven conventional result printouts.

Prior to taking part in the study, none of the participants had ever seen Visual Clot technology. However, the number of correct decisions, time to decision, diagnostic confidence, and workload improved greatly with the new tool compared with conventional ROTEM result presentation [15]. This study reveals high agreement between user perceptions and the function of the technology described in the previous Visual Clot study. This correspondence between expected and provided function is another essential feature for the success of new technologies. If this agreement does not occur, it can lead to false expectations, frustration, and ultimately reduced user acceptance. Most of the answers given concerning its design features were positive (29/36, 81%). Nevertheless, we identified some design features that led to confusion. For example, one visualization aimed to show that a test for heparin was carried out but turned out negative. This visualization intended to show an excluded heparin effect, but it confused some participants. For this reason, we modified the current version of Visual Clot. Now, it only displays heparin as present or not present. This is more in line with its situation awareness–oriented intention to show only essential information about coagulation disorders. Moreover, based on the participants' feedback, we further refined the visualization of hyperfibrinolysis and fibrin deficiency. We follow both quantitative (ie, participants' performance with a particular visualization) and qualitative study results in such design modifications. Qualitative feedback is often valuable to understand why a particular visualization is not working. This provides clues on how we can adapt it.

Eighty-six percent (36/42) of the participants agreed to the statement that Visual Clot allowed them to grasp complex coagulation situations more quickly. Ninety percent (38/42) of the participants considered it helpful to have Visual Clot and the conventional ROTEM on the screen next to each other. We regard such an integrated display mode as optimal, as Visual Clot technology converts the continuous numerical values of standard test outputs into categorical visualizations. For example, there can be too few, a healthy amount of, or too many platelets. This simplification has the advantage that information is understood quickly and easily, bearing in mind that this reduces the resolution of given data. Indeed, 47% (17/36) of the interview participants made comments regarding the missing quantification of the disorders using Visual Clot. Hence, an ideal implementation must combine the benefits of Visual Clot, which are fast and easy situation recognition, with the high precision advantage of the conventional method. A hospital may configure the visualization limits of Visual Clot according to the values of the local coagulation algorithm. This could help enforce the local standard procedure.

The main goal of Visual Clot technology is still to simplify complex viscoelastic test results in real time to facilitate the care providers' overview of the clotting situation. This study showed how potential future users perceived this technology after using it for the first time. Opinions, such as good and clear visualizations and ability to obtain a good overview of the

clotting situation in a time-efficient manner, fully support our main intention. Acutely bleeding patients benefit from faster and more accurate diagnosis and treatment. This study made us aware about the desire of our fellow physicians for such technology, from which we can draw fundamental motivation to further develop its concept.

Limitations

This study has several limitations. The opinions obtained cannot be extrapolated to a larger population because qualitative research does not assess statistical significance. Its only goal is to collect and present the broadest possible range of opinions and views. Second, we questioned the participants after their first contact with the new technology in a controlled instruction setting. Its use in the everyday clinical routine could alter the care providers' opinions. However, we consider it unlikely that Visual Clot would not be perceived as positively in clinical routine as it was in this study setting, as we observed widespread approval by anesthesiologists and intensive care physicians, who are the potential future users. Nevertheless, further research is needed to identify the strengths and weaknesses of Visual Clot technology in a clinical setting. Another limitation concerns the selection of participants. We did not include participants

randomly, but selected them according to their availability in the daily clinical routine. However, the high participation rate in the quantitative analysis reduces potential selection bias. Finally, we conducted this study in university hospitals with high standards of care in central Europe. Users' opinions may be different in other parts of the world. More research is needed in this respect as well.

Conclusions

A group of anesthesiologists and intensive care physicians regarded Visual Clot to be intuitive, easy to learn, and useful for the decision-making process in acute clotting disorders. The limitation of the visualization technology resulting from the translation of continuous measurement values into categorical values was the most frequently mentioned potential disadvantage of the technology. A split-screen implementation may be used to combine the advantages of the visualization technology and conventional technology. In this study, Visual Clot appeared to be a well-accepted decision support tool for ROTEM-based coagulation management. Further research is needed to investigate its potential in clinical practice and medical education.

Authors' Contributions

TRR, SS, JR, DRS, CBN, and DWT helped to design the study. TRR, JR, CBN, PM, CBN, and DWT helped to collect the data. TRR, SS, CBN, and DWT helped to analyze the data. TRR, SS, JR, MG, PM, KZ, DRS, CBN, and DWT helped to write the manuscript and approved the final version. TRR and SS contributed equally.

Conflicts of Interest

The University of Zurich owns the intellectual property rights to the technology described in this manuscript and registered "Visual Clot" as a trademark. The University of Zurich and Instrumentation Laboratory Company/Werfen Corporation signed a letter of intent regarding a proposed joint development and licensing agreement to develop a product based on the concept of Visual Clot. As designated inventors, DRS, CBN, and DWT may receive royalties in the event of commercialization. PM received research grants from Pfizer and Dr. F. Köhler Chemie GmbH for an investigator-initiated trial; honoraria for scientific lectures from Abbott GmbH and Co KG, Aesculap Academy, B. Braun Melsungen, Biotest AG, Vifor Pharma, Ferring, CSL Behring, German Red Cross/Institute of Transfusion Medicine, HCCM Consulting GmbH, Heinen and Löwenstein, Pharmacosmos, and Siemens Healthcare; and prizes from Aktionsbündnis Patientensicherheit, European Society of Anaesthesiology, Lohfert-Stiftung AG, Masimo–Patient Safety Foundation, and MSD-Gesundheitspreis. KZ received honoraria or travel support for consulting or lecturing from the following companies: Abbott GmbH and Co KG, AesculapAkademie GmbH, AQAI GmbH, AstellasPharma GmbH, AstraZeneca GmbH, Aventis Pharma GmbH, B. Braun Melsungen AG, Baxter Deutschland GmbH, Biosyn GmbH, Biotest AG, Bristol-Myers Squibb GmbH, CSL Behring GmbH, Dr. F. Köhler Chemie GmbH, Dräger Medical GmbH, Essex Pharma GmbH, Fresenius Kabi GmbH, Fresenius Medical Care, Gambro Hospal GmbH, Gilead, GlaxoSmithKline GmbH, Grünenthal GmbH, Hamilton Medical AG, HCCM Consulting GmbH, Heinen+Löwenstein GmbH, Janssen-Cilag GmbH, med Update GmbH, Medivance EU BV, MSD Sharp and Dohme GmbH, Novartis Pharma GmbH, Novo Nordisk Pharma GmbH, P. J. Dahlhausen and Co GmbH, Pfizer Pharma GmbH, Pulsion Medical Systems SE, Siemens Healthcare, Teflex Medical GmbH, Teva GmbH, TopMedMedizintechnik GmbH, Verathon Medical, and ViforPharma GmbH. KZ's department receives unrestricted educational grants from B. Braun Melsungen AG, Fresenius Kabi GmbH, CSL Behring GmbH, and Vifor Pharma GmbH. No other external funding or competing interests are declared. DRS's academic department receives grant support from Swiss National Science Foundation, Swiss Society of Anesthesiology and Reanimation, Swiss Foundation for Anesthesia Research, and Vifor SA. DS is co-chair of the ABC-Trauma Faculty, sponsored by unrestricted educational grants from Novo Nordisk Health Care AG, CSL Behring GmbH, LFB Biomedicaments, and Octapharma AG. DRS received honoraria/travel support for consulting or

Multimedia Appendix 1
Visual Clot educational video.
[MP4 File (MP4 Video), 13215 KB - games_v8i4e19036_app1.mp4]

Multimedia Appendix 2
Translated field notes of the participant interviews.
[PDF File (Adobe PDF File), 246 KB - games_v8i4e19036_app2.pdf]

Multimedia Appendix 3
Translated online survey announcement.
[PDF File (Adobe PDF File), 79 KB - games_v8i4e19036_app3.pdf]

References

1. Serraino GF, Murphy GJ. Routine use of viscoelastic blood tests for diagnosis and treatment of coagulopathic bleeding in cardiac surgery: updated systematic review and meta-analysis. Br J Anaesth 2017 Jun 01;118(6):823-833 [FREE Full text] [doi: 10.1093/bja/aex100] [Medline: 28475665]

2. Schöchl H, Nienaber U, Hofer G, Voelckel W, Jambor C, Scharbert G, et al. Goal-directed coagulation management of major trauma patients using thromboelastometry (ROTEM)-guided administration of fibrinogen concentrate and prothrombin complex concentrate. Crit Care 2010;14(2):R55 [FREE Full text] [doi: 10.1186/cc8948] [Medline: 20374650]

3. Gonzalez E, Moore EE, Moore HB, Chapman MP, Chin TL, Ghasabyan A, et al. Goal-directed Hemostatic Resuscitation of Trauma-induced Coagulopathy: A Pragmatic Randomized Clinical Trial Comparing a Viscoelastic Assay to Conventional Coagulation Assays. Ann Surg 2016 Jun;263(6):1051-1059 [FREE Full text] [doi: 10.1097/SLA.0000000000001608] [Medline: 26720428]

4. Davenport R, Manson J, De'Ath H, Platton S, Coates A, Allard S, et al. Functional definition and characterization of acute traumatic coagulopathy. Crit Care Med 2011 Dec;39(12):2652-2658 [FREE Full text] [doi: 10.1097/CCM.0b013e3182281af5] [Medline: 21765358]

5. Wikkelsø A, Wetterslev J, Møller AM, Afshari A. Thromboelastography (TEG) or thromboelastometry (ROTEM) to monitor haemostatic treatment versus usual care in adults or children with bleeding. Cochrane Database Syst Rev 2016 Aug 22(8):CD007871 [FREE Full text] [doi: 10.1002/14651858.CD007871.pub3] [Medline: 27552162]

6. Schöchl H, Nienaber U, Maegele M, Hochleitner G, Primavesi F, Steitz B, et al. Transfusion in trauma: thromboelastometry-guided coagulation factor concentrate-based therapy versus standard fresh frozen plasma-based therapy. Crit Care 2011;15(2):R83 [FREE Full text] [doi: 10.1186/cc10078] [Medline: 21375741]

7. Nystrup KB, Windeløv NA, Thomsen AB, Johansson PI. Reduced clot strength upon admission, evaluated by thrombelastography (TEG), in trauma patients is independently associated with increased 30-day mortality. Scand J Trauma Resusc Emerg Med 2011 Sep 28;19:52 [FREE Full text] [doi: 10.1186/1757-7241-19-52] [Medline: 21955460]

8. Stein P, Kaserer A, Spahn GH, Spahn DR. Point-of-Care Coagulation Monitoring in Trauma Patients. Semin Thromb Hemost 2017 Jun;43(4):367-374. [doi: 10.1055/s-0037-1598062] [Medline: 28297730]

9. Hunt H, Stanworth S, Curry N, Woolley T, Cooper C, Ukoumunne O, et al. Thromboelastography (TEG) and rotational thromboelastometry (ROTEM) for trauma induced coagulopathy in adult trauma patients with bleeding. Cochrane Database Syst Rev 2015 Feb 16(2):CD010438 [FREE Full text] [doi: 10.1002/14651858.CD010438.pub2] [Medline: 25686465]

10. Spahn DR, Bouillon B, Cerny V, Duranteau J, Filipescu D, Hunt BJ, et al. The European guideline on management of major bleeding and coagulopathy following trauma: fifth edition. Crit Care 2019 Mar 27;23(1):98 [FREE Full text] [doi: 10.1186/s13054-019-2347-3] [Medline: 30917843]

11. Görlinger K, Dirkmann D, Hanke AA, Kamler M, Kottenberg E, Thielmann M, et al. First-line therapy with coagulation factor concentrates combined with point-of-care coagulation testing is associated with decreased allogeneic blood transfusion

in cardiovascular surgery: a retrospective, single-center cohort study. Anesthesiology 2011 Dec;115(6):1179-1191 [FREE Full text] [doi: 10.1097/ALN.0b013e31823497dd] [Medline: 21970887]

12. Weber CF, Görlinger K, Meininger D, Herrmann E, Bingold T, Moritz A, et al. Point-of-care testing: a prospective, randomized clinical trial of efficacy in coagulopathic cardiac surgery patients. Anesthesiology 2012 Sep;117(3):531-547 [FREE Full text] [doi: 10.1097/ALN.0b013e318264c644] [Medline: 22914710]

13. Mallett SV. Clinical Utility of Viscoelastic Tests of Coagulation (TEG/ROTEM) in Patients with Liver Disease and during Liver Transplantation. Semin Thromb Hemost 2015 Jul;41(5):527-537. [doi: 10.1055/s-0035-1550434] [Medline: 26049072]

14. Kvint S, Schuster J, Kumar MA. Neurosurgical applications of viscoelastic hemostatic assays. Neurosurg Focus 2017 Nov;43(5):E9. [doi: 10.3171/2017.8.FOCUS17447] [Medline: 29088950]

15. Rössler J, Meybohm P, Spahn DR, Zacharowski K, Braun J, Nöthiger CB, et al. Improving decision making through presentation of viscoelastic tests as a 3D animated blood clot: the Visual Clot. Anaesthesia 2020 Aug 06;75(8):1059-1069. [doi: 10.1111/anae.14985] [Medline: 32030729]

16. Endsley M. Designing for Situation Awareness: An Approach to User-Centered Design, Second Edition. Boca Raton: CRC Press, Inc; 2011.

17. Symon G, Cassell C. Qualitative Organizational Research: Core Methods and Current Challenges. Thousand Oaks, CA: SAGE Publications; 2012.

18. DeepL. URL: https://www.deepl.com [accessed 2020-11-17]

19. Wordle. URL: http://www.wordle.net/ [accessed 2020-11-17]

20. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

21. Miles M, Huberman A. Qualitative Data Analysis: A Methods Sourcebook. Thousand Oaks, CA: SAGE Publications; 1994.

22. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]

23. Gulland A. Shortage of health workers is set to double, says WHO. BMJ 2013 Nov 12;347:f6804. [doi: 10.1136/bmj.f6804] [Medline: 24222676]

24. Zimmerman JE, Kramer AA, Knaus WA. Changes in hospital mortality for United States intensive care unit admissions from 1988 to 2012. Crit Care 2013 Apr 27;17(2):R81 [FREE Full text] [doi: 10.1186/cc12695] [Medline: 23622086]

25. Tscholl DW, Rössler J, Said S, Kaserer A, Spahn DR, Nöthiger CB. Situation Awareness-Oriented Patient Monitoring with Visual Patient Technology: A Qualitative Review of the Primary Research. Sensors (Basel) 2020 Apr 09;20(7):2112 [FREE Full text] [doi: 10.3390/s20072112] [Medline: 32283625]

26. Tscholl DW, Weiss M, Handschin L, Spahn DR, Nöthiger CB. User perceptions of avatar-based patient monitoring: a mixed qualitative and quantitative study. BMC Anesthesiol 2018 Dec 11;18(1):188 [FREE Full text] [doi: 10.1186/s12871-018-0650-1] [Medline: 30537934]

27. Tscholl DW, Handschin L, Neubauer P, Weiss M, Seifert B, Spahn DR, et al. Using an animated patient avatar to improve perception of vital sign information by anaesthesia professionals. Br J Anaesth 2018 Sep;121(3):662-671 [FREE Full text] [doi: 10.1016/j.bja.2018.04.024] [Medline: 30115265]

28. Garot O, Rössler J, Pfarr J, Ganter MT, Spahn DR, Nöthiger CB, et al. Avatar-based versus conventional vital sign display in a central monitor for monitoring multiple patients: a multicenter computer-based laboratory study. BMC Med Inform Decis Mak 2020 Feb 10;20(1):26 [FREE Full text] [doi: 10.1186/s12911-020-1032-4] [Medline: 32041584]

29. Pfarr J, Ganter MT, Spahn DR, Noethiger CB, Tscholl DW. Avatar-Based Patient Monitoring With Peripheral Vision: A Multicenter Comparative Eye-Tracking Study. J Med Internet Res 2019 Jul 17;21(7):e13041 [FREE Full text] [doi: 10.2196/13041] [Medline: 31317870]

30. Tscholl DW, Rössler J, Handschin L, Seifert B, Spahn DR, Nöthiger CB. The Mechanisms Responsible for Improved Information Transfer in Avatar-Based Patient Monitoring: Multicenter Comparative Eye-Tracking Study. J Med Internet Res 2020 Mar 16;22(3):e15070 [FREE Full text] [doi: 10.2196/15070] [Medline: 32175913]

31. Blackler A, Popovic V, Mahar D. Investigating users' intuitive interaction with complex artefacts. Appl Ergon 2010 Jan;41(1):72-92. [doi: 10.1016/j.apergo.2009.04.010] [Medline: 19586618]

32. Agarwal R, Karahanna E. Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage. MIS Quarterly 2000 Dec;24(4):665. [doi: 10.2307/3250951]

33. Wachter SB, Agutter J, Syroid N, Drews F, Weinger MB, Westenskow D. The employment of an iterative design process to develop a pulmonary graphical display. J Am Med Inform Assoc 2003;10(4):363-372 [FREE Full text] [doi: 10.1197/jamia.M1207] [Medline: 12668693]

34. Drews FA, Westenskow DR. The right picture is worth a thousand numbers: data displays in anesthesia. Hum Factors 2006;48(1):59-71. [doi: 10.1518/001872006776412270] [Medline: 16696257]

35. Wittgenstein L. Tractatus logico-philosophicus. London: COSIMO CLASSICS; 2007.

36. Degani A, Jorgensen C, Iverson D, Shafto M, Olson L. On Organization of Information: Approach and Early Work. Washington, DC: National Aeronautics and Space Administration; 2009.

XSL•FO
RenderX

Abbreviations

ROTEM: rotational thromboelastometry
UKF: University Hospital Frankfurt
USZ: University Hospital Zurich

XSL•FO

RenderX

Original Paper

Choice of Leisure Activities by Adolescents and Adults With Internet Gaming Disorder: Development and Feasibility Study of a Virtual Reality Program

Narae Lee[1,2], MA; Jae-Jin Kim[1,2,3], MD, PhD; Yu-Bin Shin[1,3], PhD; Hyojung Eom[1,3], MA; Min-Kyeong Kim[1], MD, PhD; Sunghyon Kyeong[1], PhD; Young Hoon Jung[1,3], MA; Sarang Min[1], BA; Joon Hee Kwon[1], BA; Eunjoo Kim[1,2], MD, PhD

[1]Department of Psychiatry & Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea

[2]Department of Psychiatry, Yonsei University Gangnam Severance Hospital, Seoul, Republic of Korea

[3]Brain Korea 21 PLUS Project for Medical Science, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Eunjoo Kim, MD, PhD
Department of Psychiatry
Yonsei University Gangnam Severance Hospital
Eonju-ro 211, Gangnam-gu
Seoul, 06273
Republic of Korea
Phone: 82 2 2019 3345
Email: ejkim96@yuhs.ac

Abstract

Background: Excessive internet game use frequently leads to various physical, psychological, and social problems, and internet gaming disorder (IGD) has become a serious public health issue worldwide. Recently, virtual reality (VR) therapy has emerged as a promising method to increase psychological treatment motivation and accessibility. However, few studies have examined the potential of VR technology for the management of IGD, and VR content tailored to IGD characteristics remains scarce.

Objective: This preliminary study aimed to examine the potential of a VR-based program that was designed to help users identify their leisure time use patterns, especially those related to gaming, and to modify their gaming overuse by alternative activities provided in the VR content. Moreover, to investigate whether users' VR activities reflect various clinical variables of IGD in youth, we examined the relationships among the leisure time activity selection pattern, built-in response, and speech data obtained from the VR program, as well as symptom severity of internet gaming, psychiatric comorbidities, and motivation of participants reported through relevant questionnaire data.

Methods: Three types of VR content (understanding my daily activities at home, finding an alternative activity to internet gaming at home, expressing contradictory opinions toward a friend's gaming beliefs) were developed by simulating the daily situations in which patients with IGD can select alternative free-time leisure activities. We examined internet addiction, mental health problems, and motivation for 23 IGD and 29 control participants. Behavioral and self-rated responses from VR, such as alternative activity selection data and speech patterns (speech time, speech satisfaction, and speech accordance), and results from various questionnaires were compared between groups. The correlations between IGD behaviors in VR and real-life behaviors assessed by questionnaire measures were analyzed.

Results: Significant correlations were found between internet gaming behavior and user activity data, such as speech and activity selection pattern, in our VR program. Our results showed that the IGD group had fewer leisure activities and preferred game or digital activities to other types of activities compared to controls, even in VR. There was a positive relationship between the viability of alternative leisure activities the participants selected in VR and the amount of perceived satisfaction from that activity (r=.748, *P*<.001). Speech accordance in the IGD group was lower than in the control group and was correlated negatively with Internet Addiction Test and Internet Addiction Test–gaming scores (r=.300, *P*=.03) but positively with users' motivation (r=.312, *P*=.02).

Conclusions: The results from our VR program can provide information about daily activity patterns of youths with IGD and the relationship between user VR activities and IGD symptoms, which can be useful in applying VR technology to IGD management.

XSL•FO

RenderX

Introduction

Internet gaming is the most common leisure activity among adolescents in East Asian countries, including South Korea [1,2] and China [3]. While some studies have proposed positive effects of internet gaming [4,5], excessive internet game use frequently leads to various physical, psychological, and social problems [6-8]. The concept of internet gaming disorder (IGD), defined as the persistent and recurrent use of the internet to engage in games despite negative consequences, has been controversial. Diagnostic criteria or guidelines for IGD have been proposed by several investigators, but there has been continued debate around the conceptualization of IGD. Therefore, studies of IGD-targeted treatment mechanisms are needed to help with the classification of behavioral addiction and the effective management of individuals with IGD [9-11].

Regarding IGD treatment, theories based on cognitive behavioral therapy (CBT) and motivational enhancement therapy (MET) have been promising approaches to explain addictive behaviors [12], and they have been commonly used as effective treatment for IGD [12-14]. The CBT model claims that the identification and modification of maladaptive cognition on internet gaming are crucial factors for treatment [14]. Maladaptive cognitions such as excessive preoccupation with internet gaming, impaired cognitive control, and cognitive inflexibility contribute to compulsive internet gaming activity [13,15,16].

Additionally, the motivation-focused model of addiction has been employed for the management of IGD, as motivational drives linked to reward-seeking also contribute to problematic gaming behavior [15,17]. Individuals with IGD are reported to have enhanced reward sensitivity and decreased loss sensitivity, leading to stronger motivation to play compared to people without IGD [15,16,18].

Taken together, CBT and MET models of addiction suggest that treatment for IGD requires assessing the individual's cognitive and behavioral patterns and motivation before a specific therapeutic plan is implemented. Therefore, the management of the selection process for gaming activities is important for finding the triggering factors that motivate and maintain people's excessive gaming behaviors [6,19]. Since people with IGD often do not have insight on their excessive preoccupation with gaming, it is necessary for patients with IGD to identify the pattern and underlying motivation of their leisure activities to manage their problematic gaming effectively. For example, identification of the times and places that elicit people's craving for gaming in their free time and a technique to control this craving and the motivational components of internet gaming is necessary for IGD management [20]. For this purpose, therapists also need to observe and intervene in the behavior of individuals with IGD in response to exposure cues for leisure activities in situations that increase the patients' risk of gaming. However, it is difficult to reproduce the actual reality of the gaming cue exposure through the process of traditional treatment. Besides, in a real-world situation, it is also difficult to follow up and collect patients' behavioral data in real time.

Alternatively, CBT using virtual reality (VR) technology is a promising method to increase motivation for and accessibility of mental health treatments, especially among youth who are generally positive to the appeal of modern digital technology [12-14,21-23]. Moreover, VR features are immersive and interactive, simulating reality without space and time limits [24]. Gestures and voice inputs are effective means of interacting with avatars in virtual space while also maintaining full potential for flow and immersion [24,25]. The VR program also automatically records users' activity and verbal data, allowing for real-time data analysis. Despite these advantages of VR-CBT, few studies have examined the effects of VR for IGD management, and VR content tailored to IGD characteristics is scarce.

Regarding the embodiment of IGD management techniques, we mainly used the self-speech technique. The definition of "self-speech" in our program was the verbalization of one's opinions in their own words in response to the questions asked by avatars. Self-speech is reported to help to self-regulate cognition and behavior in addition to positively influencing self-control [26,27]. It also helps people accomplish self-distancing and objectification, enabling individuals with IGD to reappraise their cognition on gaming guided by objective viewpoints in the VR program [28]. As computer technologies related to voice recognition and processing have developed, speech has also become a viable interaction modality in VR environments.

Given this background, the objectives of this study were to examine the potential of a VR-based program that was designed to help users identify their leisure time use patterns, especially those related to gaming, and to modify their gaming overuse by alternative activities provided in the VR content. More specifically, to investigate whether behavioral responses to our VR program were associated with various IGD clinical variables, we examined the relationships among leisure time activity selection pattern, built-in response and speech data (speech time, speech satisfaction, and speech accordance) obtained from our VR program, and IGD symptom severity, psychiatric comorbidities, and motivation of participants reported through relevant questionnaire data.

Methods

Sample Size (Power)

Power analysis for an independent t test was conducted in the G*Power program to determine the necessary sample size, with an alpha of .05, power of 0.80, large effect size (f =0.8), and 2 tails [29]. The results from our sample size calculation suggested that the desired sample size was estimated to be 26 in each

group, totaling 52 subjects, to detect the group difference in our dependent measures.

Recruitment

Participants were recruited from online community and social network sites in South Korea and through flyers. Since men are known to have a higher prevalence of IGD than women, and to avoid gender-specific confounding factors related to the pattern of gaming craving and motivations affecting the results, only male participants were recruited [30]. Participants were 52 men aged 11-25 years, and 44 (44/52, 85%) were either high school or university students, aged between 16 and 25 years. Age was matched between the 2 groups during the assignment procedure, but other variables in our study were the result of simple randomization. All participants were evaluated for internet use patterns and interviewed by a psychiatrist for IGD diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [31]. In the control group, no one had ever been diagnosed with IGD or other DSM-5 psychiatric disorders. Exclusion criteria for all participants were current use of psychotropic medication and history of substance use disorder, serious neurological or medical disorder, bipolar I disorder, and psychotic disorder according to the Mini-International Neuropsychiatric Interview for Children and Adolescents (MINI-KID 6.0) [32] for participants younger than 18 years and the MINI 5.0 for adult participants [33]. However, 3 participants with psychiatric comorbidities were included (1 attention deficit hyperactivity disorder [ADHD], 1 depressive disorder, 1 adjustment disorder) in the IGD group.

Measures

IQ was assessed using the short form of the Wechsler Intelligence Scale for Children—Third Edition [34] and Wechsler Adult Intelligence Scale—Revised [35,36]. The Internet Addiction Test (IAT) [8], one of the most utilized diagnostic instruments for internet addiction, was used to assess IGD symptom severity (Cronbach α=.90). All participants completed a modified version of the IAT that replaced the term "Internet" with terms such as "Internet games" to assess specific subtypes of internet addiction [37-39]. As IGD is frequently comorbid with other psychiatric disorders [40], participants completed the short version of Conners-Wells' Adolescent Self-Report Scale (CASS-S) for adolescents (Cronbach α=.88) [41] and Conner's Adult ADHD Rating Scale (Cronbach α=.87) [42] for adult participants to measure ADHD symptoms [42,43], Center for Epidemiological Studies Depression Scale (CES-D) for depression (Cronbach α=.91) [44,45], and Beck Anxiety Inventory (BAI) for anxiety (Cronbach α=.92) [46]. The Free Time Motivation Scale for Adolescents (FTMS-A; Cronbach α=.75) was also administered to measure 5 subscales representing various types of motivation to engage in certain types of leisure time activities [47].

The presence questionnaire (PQ; Cronbach α=.84) [48,49], Simulator Sickness Questionnaire (SSQ) [50], and Client Satisfaction Questionnaire (CSQ) [51] were used to measure the users' experience with our system. PQ is a measure of users' presence in virtual reality, which is a psychological state of "being there" and an important element of VR experience. Presence can enhance users' active engagement in content,

involving their senses and capturing their attention. SSQ assesses simulator sickness resulting from the discrepancy between simulated visual motion and the sense of movement coming from the vestibular system. Simulator sickness is negatively correlated with users' enjoyment of VR programs. Finally, CSQ was used to rate the users' general satisfaction with our VR training session.

Procedures

A between-subjects user study was conducted for about 1 hour. Participants were asked to complete written questionnaires about their demographic background and internet usage patterns, including the average hours they spent online gaming, frequency of internet usage, and name of a game they mainly play. Participants were adjusted to the virtual environment during the training session and then were asked to play 3 types of VR content, described in the following sections. The content was counterbalanced to prevent systemic order effect. Upon completing the program, participants were debriefed about their VR experience, using the postquestionnaire form.

Virtual Reality

System Design

The VR system was a desktop computer running Microsoft Windows 10 containing an NVIDIA GeForce GTX 970 graphics card and an Oculus Rift head-mounted display with tracker (Oculus VR, LLC, Irvine, CA; HD resolution of 1080×1200 per eye with 51.6 diagonal field of view, a 3-DOF tracker for head rotation, and built-in headphones). Mounted on the headset, the Leap Motion Controller (Leap Motion Inc, San Francisco, CA), a new device suitable for hand gesture–controlled user interfaces, was used for interactions with executable objects and avatars during the VR experience (Multimedia Appendix 1). The microphone built into the VR headset gathered verbal data from users' self-speech in real time. Three-dimensional virtual environments included a virtual house and appliances using Autodesk 3D max. User instruction was provided verbally. The virtual environment was integrated with Unity software. User data such as type of selected activities and selection time, head movement, hand gesture, speech content, and speaking time were recorded and stored in the main server computer. Menu navigation on screen was performed by virtual hands positioned to match the user's hands. This system allowed the therapist to track patient performance and analyze behavioral information.

Content Design

Our VR training system emphasized the development of motivation for change based on the implementation of CBT and MET principles by: (1) identifying situations, thoughts, and behavior associated with internet game use by providing situations that increase the patients' risk of gaming and dysfunctional beliefs that support the maintenance of behavioral addiction issues; (2) conducting cognitive restructuring with these dysfunctional beliefs by describing the pros and cons of his or her online activity, expressing contradictory opinions against irrational beliefs about games; and (3) raising one's motivation by devising a self-determined choice of activities to modify one's maladaptive leisure time use pattern.

Content 1 involved "understanding my daily leisure activities at home" and was designed to understand the pattern of participants' choice of leisure activities in daily life. It provided real-life situations that increase the patients' risk of internet gaming and their dysfunctional behavioral patterns that support the maintenance of gaming issues. Figure 1A shows a blueprint menu of an indoor apartment showing the whole space for finding leisure activities. It consisted of a porch, participant's room, small room, living room, kitchen, and restroom. At the start of each level, the entrance of the apartment was displayed from the first-person perspective. When users selected a room in the House Map menu with their hands, the scene automatically started at the entrance of the chosen room. The program then suggested users explore and select a set of leisure activities by choosing virtual objects placed in each room (Figure 1B). After selecting each activity, 2 questions were presented on the head-mounted display screen. Users determined the proportion of the chosen activity during daily life measured using a visual analog scale (VAS) ranging from 0% to 100%. Missions A and B consisted of the same virtual environment, but Mission A aimed to discover leisure activities during the daytime, while Mission B targeted activities during the nighttime. Multimedia Appendix 2 explains the list of leisure activities users can select in each room.

Content 2 involved "finding an alternative activity to internet gaming at home" and was designed to encourage users to discover alternative activities the user can substitute for game playing. This is analogous to the alternative thought recording, which is one of the most commonly used CBT techniques to modify an individual's irrational cognition. We tried to figure out the behavioral differences between the IGD group and control group through the alternative activity selection pattern in VR and self-reported viability. The virtual environment of Content 2 was identical to that of Content 1. The program asked users to find leisure activities commonly found at home to replace internet gaming. Once an activity was selected, the

self-assessment for perceived viability ("How likely are you actually to do the selected activities at home in your daily life?") and perceived satisfaction ("How satisfied do you think you will be when you participate in the selected activity?") of the selected activity were measured using a VAS (Multimedia Appendix 3). Additionally, an Add Activity menu was provided to add custom activities that are not found in the apartment to maximize user autonomy. After choosing all leisure activities, visualized training feedback results that were analyzed based on the participants' own activities in VR were displayed as a pie diagram (Figure 1C). The program also automatically recorded activity type and corresponding response, allowing for real-time data analysis. The activities on a computer, mobile phone, and console game were included in the digital activity category for analysis.

Content 3 involved "expressing contradictory opinions toward friend's gaming beliefs" and was designed for users to reappraise their cognition regarding games, rebutting common motivations for playing games. In Mission A, we collected participants' self-speech data in which they express contradictory opinions against male avatar's beliefs about games and recommend alternative activities they chose in Content 2 to the avatar friends in VR. Then, users were asked to express contradictory opinions toward the avatar peers' beliefs regarding games through self-speech about 6 topics that are based on gaming motivation theory of socialization, achievement, and dissociation (see the detailed scripts of these topics in Multimedia Appendix 4). Also, the users were given a VR content wherein an avatar peer-pressured the user to play a game together to relieve stress from exams or to win a battle against their opponents. The users were instructed to refuse the avatar's invitations to play the game by stating the reason or to provide alternative activities to gaming; whether the users could refuse their peers' proposals to play the game was observed. Through this training, users could also review their gaming motivations and ensure that self-assertiveness is achieved properly.

Figure 1. Screenshots of apartment setting (A), gestural input (B), and questionnaire (VAS) and visual feedback for selected activities (C) in VR.

In Mission B, the users' own selection of alternative activities from Content 1 and Content 2 were displayed as a mind map on the TV screen, helping users to make a 1-minute speech about the strengths of alternative activities on physical, social, and psychological aspects in comparison to gaming and internet use. All the speech content was automatically recorded when users pressed a speech button during the VR program. After each speech session, participants were required to rate their speech using a VAS in terms of 3 elements: speech time (duration of speech measured in milliseconds), speech satisfaction (self-assessment of satisfaction of one's speech), and speech accordance (concordance of self-speech contents they spoke and users' actual thoughts; "How much does your opinion in reality match what you just said?").

Statistical Analysis

The main dependent variables in each group were normally distributed, as assessed by a Shapiro-Wilk test ($P>.05$). All demographic and clinical symptom measures were compared between the IGD and control groups using independent t tests

(2-tailed). Bivariate correlation analysis (Pearson r) was conducted to examine the association between behavioral and speech results on VR tasks and IGD symptom severity assessed by IAT and IAT-gaming and various clinical questionnaire measures, such as CASS-S, CES-D, BAI, and FTMS-A. All analyses were conducted using SPSS 20.0 (SPSS Inc, Chicago, IL). In all cases, $P<.05$ (2-tailed) was considered statistically significant.

Ethics

The current study was approved by the Institutional Review Board of Yonsei University College of Medicine, Gangnam Severance Hospital, and written informed consent was obtained from all participants after they had received a detailed explanation of the study. In case of adolescents, written informed consent of their parents was also obtained.

Results

Comparison of Clinical Symptom Measures Between IGD and Control Groups

Table 1 shows the result of the 2-tailed independent *t* tests between the IGD and control groups in demographic and clinical variables. There were no differences in age and IQ between groups. As expected, the IGD group had higher IAT and IAT-gaming scores. A significantly higher CASS-S score was also reported for the IGD group, suggesting that the IGD group showed higher ADHD symptoms compared to controls. Moreover, IAT and BAI (r_{52}=.314, *P*=.02), IAT and CES-D (r_{52}=−.274, *P*=.049), IAT-gaming and BAI (r_{52}=.339, *P*=.01), and IAT-gaming and CES-D (r_{52}=−.260, *P*=.06) were significantly correlated, indicating that the severity of anxiety and depression symptoms was significantly associated with IGD severity.

Table 1. Demographic and clinical characteristics of participants.

Characteristics	IGD[a] group, (n=23), mean (SD)	Control group, (n=29), mean (SD)	*t* test[b]	*P* value
Age (years)	18.91 (3.87)	19.75 (3.87)	0.788	.44
IQ	113.6 (14.37)	117.6 (13.06)	0.038	.32
IAT[c]	64.87 (15.53)	37.68 (11.58)	−6.998	.00
IAT-gaming	64.23 (15.09)	35.08 (12.76)	−8.940	.00
CES-D[d]	14.83 (9.01)	10.83 (8.07)	−1.860	.07
BAI[e]	7.09 (7.39)	3.76 (4.05)	−2.068	.44
ADHD[f] score	26.61 (14.73)	16.17 (9.03)	−3.146	.003
PQ[g]	122.87 (26.67)	133.10 (17.84)	1.654	.10
SSQ[h]	24.35 (7.64)	25.28 (10.26)	0.361	.72
CSQ[i]	11.65 (2.85)	13.31 (2.46)	2.246	.29

[a]IGD: internet gaming disorder.

[b]The degrees of freedom for all *t* tests were 51.

[c]IAT: Internet Addiction Test.

[d]CES-D: Center for Epidemiological Studies Depression Scale.

[e]BAI: Beck Anxiety Inventory.

[f]ADHD: Attention Deficit Hyperactivity Disorder.

[g]PQ: Presence Questionnaire.

[h]SSQ: Simulator Sickness Questionnaire.

[i]CSQ: Client Satisfaction Questionnaire.

Leisure Activity in VR Content and Symptoms of IGD

In Content 1, we tried to analyze the number and patterns of daily leisure activities in VR in the IGD and control groups. Our results showed that the IGD group had fewer activity domains than controls (t_{51}=3.529, *P*<.001; Figure 2A). Interestingly, even in the VR content, there were significant differences in the percentage of digital or game activities (ratio of game or digital leisure activities to total selected activities) between the IGD group and controls (t_{51}=−2.79, *P*=.007; Figures 2B and 2C). The IGD group preferred game or digital activities to other types of leisure activities more than controls did, and chosen activities were mostly in the mobile and computer category rather than being varied (Figure 2D).

In Content 2, correlation analysis showed that the number of alternative activities chosen by each participant was negatively correlated with both the IAT (r_{52}=−.341, *P*=.01) and IAT-gaming (r_{52}=−.280, *P*=.04) scores (Figure 3A). Additionally, the percentage of digital activities in VR was positively related to depression (r_{52}=.416, *P*=.002) and anxiety (r_{52}=.321, *P*=.02; Figure 3B), indicating that selected activity in VR can closely relate to psychiatric symptoms. Lastly, there was a positive relationship between the viability of alternative leisure activity participants selected in the VR environment and the amount of perceived satisfaction from that activity (r_{52}=0.748, *P*<.001).

XSL•FO

RenderX

Figure 2. Frequency of digital and game activities in the virtual reality environment in Content 1: (A) number of selected activity categories in level 1, (B) frequency of the selection of digital and gaming activities, (C) proportion of the selection of gaming activities, (D) percentage of nongaming and gaming activities in each group. IGD: internet gaming disorder.
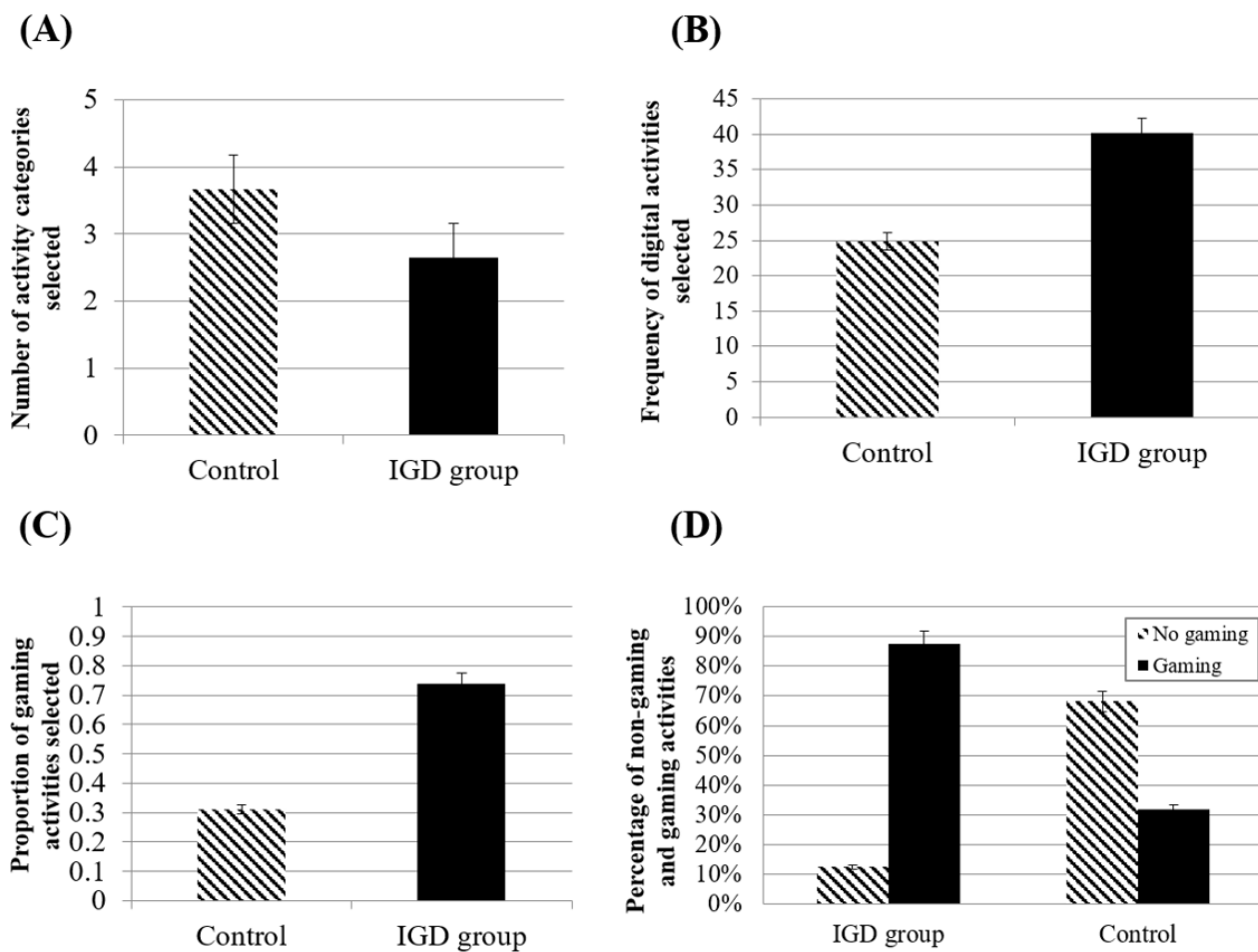
Figure 3. Relationships between the Internet Addiction Test (IAT) gaming score and (A) number of leisure activities in the virtual reality (VR) environment and (B) IAT score as well as relationships between the percentage of digital activity in the VR environment and (C) depression and (D) anxiety. BAI: Beck Anxiety Inventory; CES-D: Center for Epidemiological Studies Depression Scale.
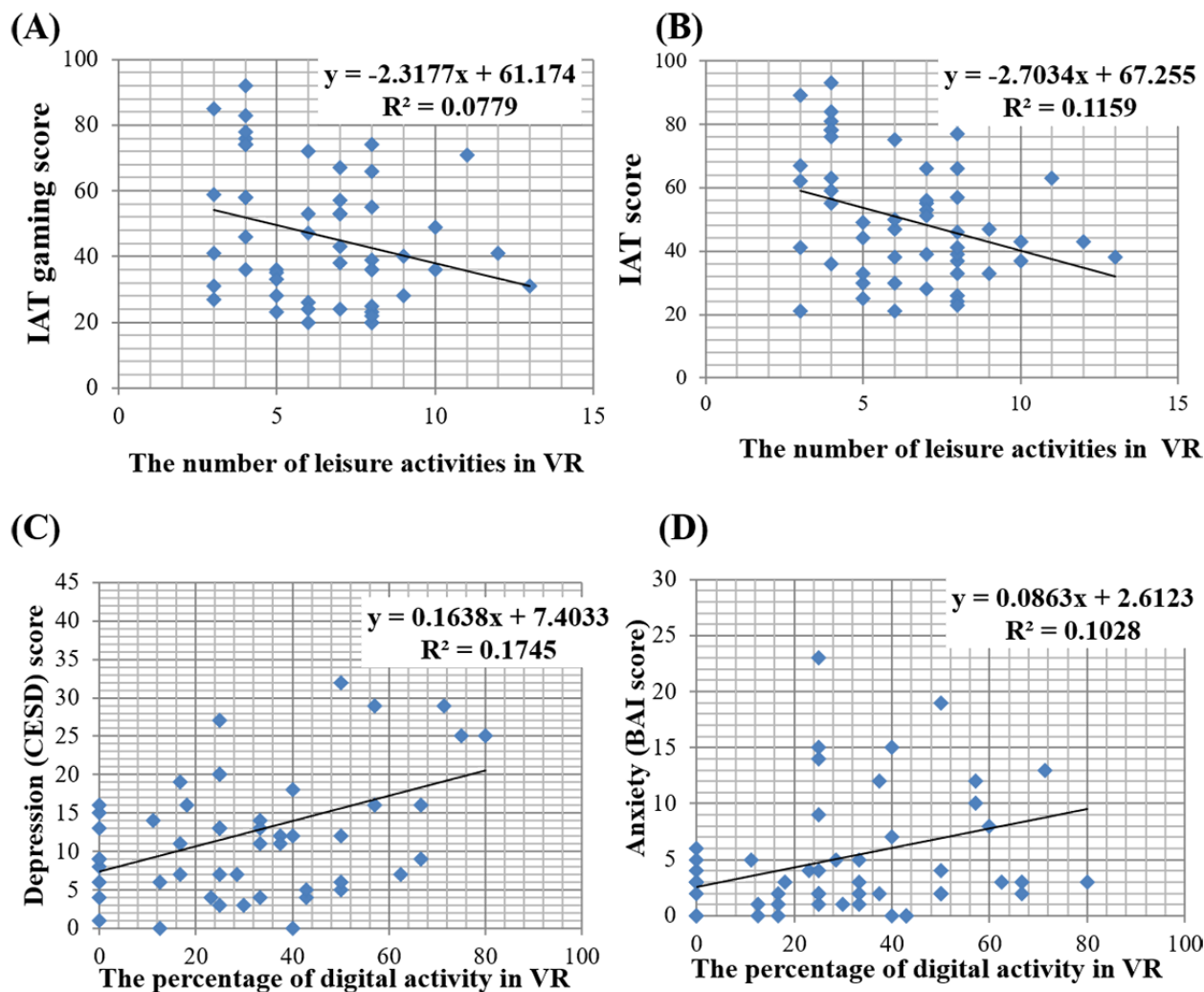
Speech Patterns About the Choice of Leisure Activities in VR

In Content 3, the independent t tests showed that users' speech time and speech satisfaction were not significantly different between the IGD and control groups, but speech accordance in the IGD group was lower than in the control group (t_{51}=2.762, P=.008; Figure 4A). Additionally, speech accordance was negatively correlated with IAT (r_{52}=−.375, P=.006; Figure 4B) and IAT-gaming scores (r_{52}=−.376, P=.006; Figure 4C), but positively correlated with motivation (r_{52}=.300, P=.03; Figure 4D).

The average speech satisfaction score in VR was positively correlated with motivation (r_{52}=.312, P=.02) and perceived pleasure of the activity (r_{52}=.31, P=.03; Figures 5A and 5B). Speech satisfaction score was negatively associated with depression (r_{52}=−.392, P=.004) and frequency of digital activity in VR content (r_{48}=.290, P=.04; Figures 5C and 5D). Independent t tests reported no significant differences in PQ, SSQ, and CSQ scores between the 2 groups. This shows that both IGD and control groups had similar experience in terms of immersion, simulator sickness, and program usability.

Figure 4. (A) Comparison of the accordance of self-speech contents and opinion between the control and internet gaming disorder (IGD) groups and correlations between accordance of self-speech and opinion and (B) Internet Addiction Test (IAT) score, (C) IAT-gaming score, and (D) Free Time Motivation Scale (FTMS)-intrinsic motivation.
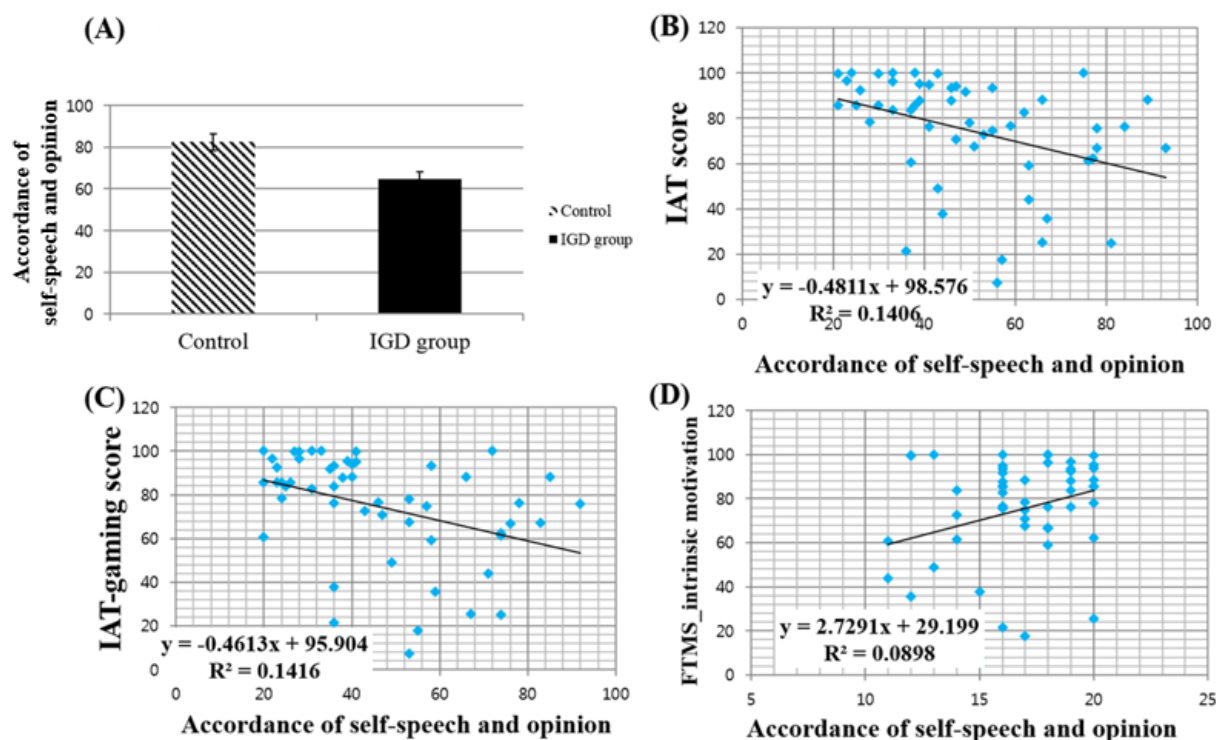
Figure 5. Correlation between speech satisfaction and (A) Free Time Motivation Scale (FTMS)-intrinsic motivation, (B) perceived pleasure, (C) depression, and (D) percentage of digital activities in virtual reality (VR). CES-D: Center for Epidemiological Studies Depression Scale.
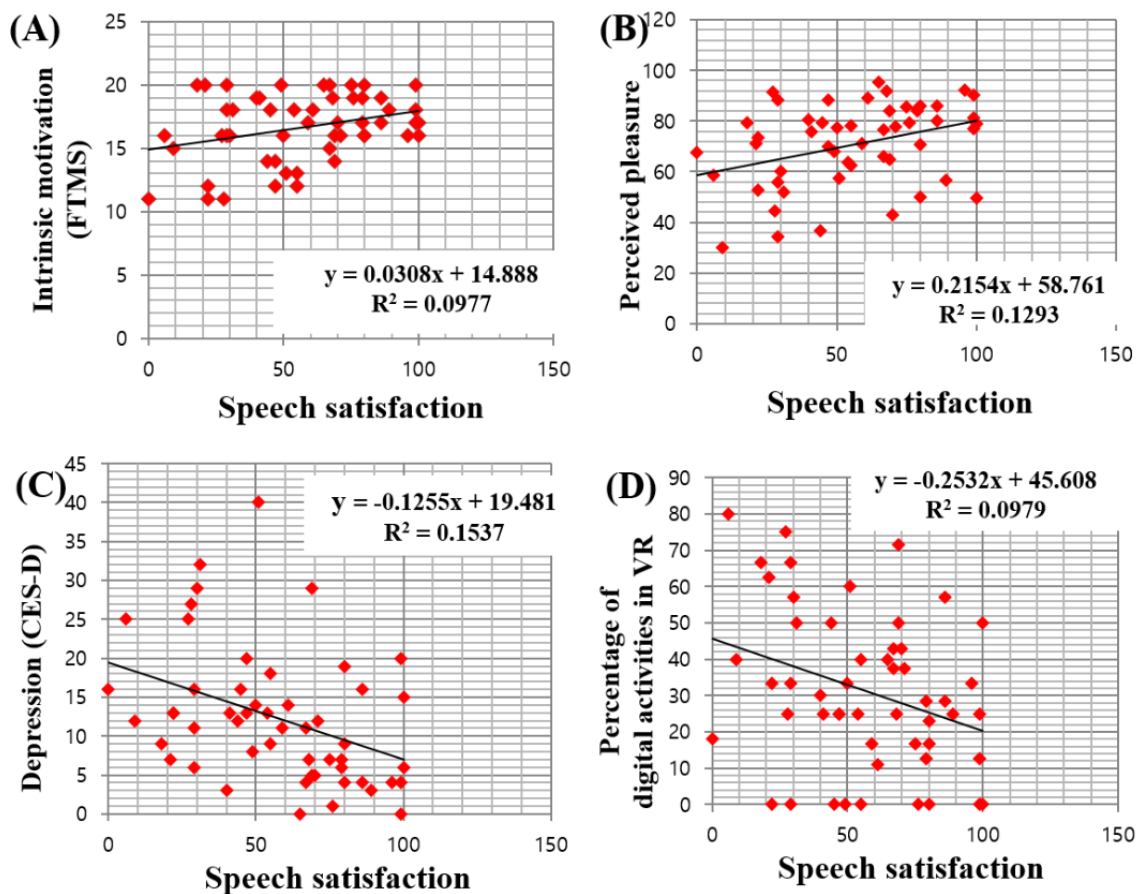
Discussion

Our VR program aimed to help IGD patients to identify their game-related behavior patterns by simulating the daily situations where they can select leisure activities and manage risky situations that lead to gaming. We also found significant relationships between the behaviors of youth with IGD in VR and in daily life by presenting the significant relationships among leisure activity selection patterns, built-in response data obtained from the VR program, and IGD symptom severity, psychiatric comorbidity, and motivation of participants reported through relevant questionnaire data. With varied VR content, users were expected to autonomously engage in self-help training with enhanced accessibility and without face-to-face supervision from therapists.

Based on CBT and MET principles, our VR program was constructed in a way that participants could practice managing malcognition about internet gaming and foster motivation and self-determination skills in an immersive virtual environment. Users were encouraged to think reflectively about their previous gaming behaviors and irrational game-related cognition. For example, one of the representative maladaptive patterns for cognition of patients with IGD who had "excessive preoccupation with internet gaming" was revealed through their overuse of digital-related activities, even in VR. Also, the real-time visual feedback report of selected activities using pie charts (Figure 1C) was intended to help patients to recognize their maladaptive leisure time behavior patterns and can potentially be useful to establish objective behavioral evidence to assess their IGD symptoms.

Moreover, the users were encouraged to increase their motivation to change by being given the opportunities for their own choice and self-direction while choosing free-time activities in our VR program. For example, the users were allowed to add their own activities in the Add Activity menu. Also, the self-rating component of perceived viability and perceived satisfaction of selected activities made it possible for the users to evaluate their motivation and acknowledge their own feelings related to this motivation. In addition, through self-speech in our VR program, users were given chances to voluntarily develop reflective thought-processing skills by considering the pros and cons of internet gaming.

The data from our VR program can provide information about daily activity patterns of youths with IGD, as well as the relationship between user VR activities and IGD symptoms. This suggests that IGD symptoms can influence the users' selection patterns for leisure activities and ability to discover alternative activities to gaming. The finding that the IGD group had fewer leisure activities and preferred game or digital activities to other types of activities, even in VR, may reflect the lack of interest or pool of leisure activities in youth with IGD beyond games [15,16]. This was consistent with previous studies reporting cognitive bias (eg, attention and approach) to addiction-related stimuli in addiction-prone individuals [52]. Similarly, patients with IGD are known to have a general preoccupation with internet gaming and overvaluation of game-related activities and reward [15,16,53]. Our result was consistent with motivation-focused models postulating that addiction may be a disorder of misdirected motivation, in which greater priority is given to addiction-related stimuli [15,16,18].

Additionally, there was a positive correlation between the viability of alternative leisure activities participants selected and the amount of perceived satisfaction of the activities or user's motivation. This suggests that it is important to train youths with IGD to recognize and engage in various types of alternative activities, rather than indulging only in internet gaming, during their leisure time. Through our VR program, individuals with IGD may broaden their range of leisure activities by getting more exposure to new alternatives to replace internet gaming. This finding also shows the potential of our VR-based training to predict and prevent IGD-related symptoms, by providing the opportunity for people to observe their behavior patterns in VR and practice selecting alternative activities of their own choice.

For IGD treatment, VR therapy is efficacious in decreasing craving and severity of internet addiction [54,55]. However, unlike VR therapy for substance addictions or pathological gambling, few studies have formally examined the efficacy of VR therapy for IGD, and evidence-based treatment components have not been well-established [13,56]. To base our program on clinically proven evidence, our program tried to incorporate the treatment component of traditional CBT and MET models of addiction [14,57]. Furthermore, adopting the motivational aspect of addiction, our results, such as the positive correlation between speech satisfaction or accordance and users' motivation, negative correlation between speech satisfaction and the percentage of digital activities in the VR content, the IAT and IAT-gaming scores, and depressive symptoms indicating the behavior pattern, speech satisfaction, and speech accordance measured in our VR program, may be useful to assess motivation, IGD symptom severity. and affective aspects of youth [58].

Regarding the results of the PQ, SSQ, and CSQ, the IGD and control groups had similar levels of satisfaction in terms of immersion, simulator sickness, and program usability.

Although our findings showed that our VR may predict IGD-related symptoms through VR-based user activities, this study had several limitations to consider. The small sample size, inclusion of only male participants, and relatively mild level of IGD severity among our participants limit the generalizability of our findings to female populations and those with severe symptoms. Also, participants of a diverse age range were recruited for this study. Children, adolescents, and young adults may have their own unique characteristics in gaming motivation and various leisure activities. Therefore, a detailed subgroup analysis regarding the difference between age groups is needed in future studies. Moreover, 3 patients with comorbid psychiatric disorders were included, for which the presence of a comorbid condition may be a confounding factor. However, the main findings of our study generally remained the same as those when including participants with comorbidities. Additionally, symptoms of IGD or other psychopathologies were measured by self-report instruments, and objective behavioral observation of users or reports from caregivers are required. Also, the

FTMS-A, originally developed for the adolescent population, was used for adult participants for the ease of comparison between adolescent and young adult groups; however, the use of a motivation scale for the adult population is still needed. Moreover, the measured variables of self-speech were limited to speech time or users' subjective self-rated variables. For deeper qualitative analysis, meaning-based speech content analysis is required. Furthermore, with the current technical limitations, previously reported simulator sickness [50,59-61] made it difficult for people to experience VR content that exploit user motion or the implementation of a moving environment in VR. In addition, the avatar cannot give any verbal or nonverbal feedback to the users or afford any other form of interactivity; thus, our program was limited in terms of responsiveness and realism in the virtual interaction landscape. With future technical advancement in motion capture, face capture, speech recognition, and acoustic, meaning-based speech analysis, inputs from VR systems will be fed into an artificial intelligence model that can determine the user's goals and provide appropriate outputs that can be translated into speech and gestures of avatars.

Also, some discrepancy may exist between leisure activities chosen in the VR and those in real life, as shown by the low speech accordance rate in IGD participants. Although our experiment was conducted in a controlled hospital setting with assistance from experimenters, further research is needed in the home environment using mobile or wearable devices to investigate the potential of our program as a standalone self-help therapy tool. Additionally, the role of gaming motivation as a mediating factor in the development of IGD was not fully explored in this study. It would be informative to investigate whether different motivations for playing internet games mediate or moderate the relationship between anxiety or depressive symptoms and IGD severity. Lastly, our preliminary study design did not allow for validation of our program to reduce IGD symptoms or related outcomes, such as the time spent on gaming or depressive symptoms, and various reliability measures, such as test-retest reliability, were also not measured. In future studies, different types of validity testing should be conducted with larger sample sizes to validate the clinical usefulness of our program.

Our results showed significant correlations between behaviors of adolescent and young adults with IGD in VR and in daily life, in terms of the selection and verbalization of their leisure activities and related cognition. If a database is accumulated for more precise speech content analysis with a larger sample size, this system can potentially serve as a useful tool for researchers to understand IGD more deeply and also have a considerable impact on designing interaction and interface in VR content for patients with IGD.

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1
Virtual reality system and experimental environment.
[PNG File , 419 KB - games_v8i4e18473_app1.png]

Multimedia Appendix 2
Leisure activities in each room in Contents 1 and 2.
[DOCX File , 13 KB - games_v8i4e18473_app2.docx]

Multimedia Appendix 3
Screenshots for Content 2: (A) friend avatar's opinion; (B) self-speech; (C) self-report on accordance between speech and their opinion.
[PNG File , 309 KB - games_v8i4e18473_app3.png]

Multimedia Appendix 4
Script of Content 3a based on 3 gaming motivations.
[DOCX File , 14 KB - games_v8i4e18473_app4.docx]

References

1.   Jeong EJ, Kim DH. Social activities, self-efficacy, game attitudes, and game addiction. Cyberpsychol Behav Soc Netw 2011 Apr;14(4):213-221. [doi: 10.1089/cyber.2009.0289] [Medline: 21067285]

XSL•FO

RenderX

2.   Kim K, Ryu E, Chon M, Yeun E, Choi S, Seo J, et al. Internet addiction in Korean adolescents and its relation to depression and suicidal ideation: a questionnaire survey. Int J Nurs Stud 2006 Feb;43(2):185-192. [doi: 10.1016/j.ijnurstu.2005.02.005] [Medline: 16427966]

3.   von Deneen KM, Yuan K, Qin W, Tian J, Liu Y, Gold MS. Internet Addiction. Directions in Psychiatry 2014;34(2):119-127.

4.   Durkin K, Barber B. Not so doomed: computer game play and positive adolescent development. Journal of Applied Developmental Psychology 2002 Nov;23(4):373-392. [doi: 10.1016/S0193-3973(02)00124-7]

5.   Gamberini L, Barresi G, Maier A, Scarpetta F. A game a day keeps the doctor away: A short review of computer games in mental healthcare. Journal of CyberTherapy and Rehabilitation 2008;1(2):127.

6.   Allison SE, von Wahlde L, Shockley T, Gabbard GO. The development of the self in the era of the internet and role-playing fantasy games. Am J Psychiatry 2006 Mar;163(3):381-385. [doi: 10.1176/appi.ajp.163.3.381] [Medline: 16513856]

7.   Davis R. A cognitive-behavioral model of pathological Internet use. Computers in Human Behavior 2001 Mar;17(2):187-195. [doi: 10.1016/S0747-5632(00)00041-8]

8.   Young KS, Rogers RC. The Relationship Between Depression and Internet Addiction. CyberPsychology & Behavior 1998 Jan;1(1):25-28. [doi: 10.1089/cpb.1998.1.25]

9.   Rumpf H, Achab S, Billieux J, Bowden-Jones H, Carragher N, Demetrovics Z, et al. Including gaming disorder in the ICD-11: The need to do so from a clinical and public health perspective. J Behav Addict 2018 Sep 01;7(3):556-561 [FREE Full text] [doi: 10.1556/2006.7.2018.59] [Medline: 30010410]

10. Saunders JB, Hao W, Long J, King DL, Mann K, Fauth-Bühler M, et al. Gaming disorder: Its delineation as an important condition for diagnosis, management, and prevention. J Behav Addict 2017 Sep 01;6(3):271-279 [FREE Full text] [doi: 10.1556/2006.6.2017.039] [Medline: 28816494]

11. Chen KH, Oliffe JL, Kelly MT. Internet Gaming Disorder: An Emergent Health Issue for Men. Am J Mens Health 2018 Jul;12(4):1151-1159 [FREE Full text] [doi: 10.1177/1557988318766950] [Medline: 29606034]

12. Young KS. Cognitive behavior therapy with Internet addicts: treatment outcomes and implications. Cyberpsychol Behav 2007 Oct;10(5):671-679. [doi: 10.1089/cpb.2007.9971] [Medline: 17927535]

13. Dong G, Potenza MN. A cognitive-behavioral model of Internet gaming disorder: theoretical underpinnings and clinical implications. J Psychiatr Res 2014 Nov;58:7-11 [FREE Full text] [doi: 10.1016/j.jpsychires.2014.07.005] [Medline: 25062755]

14. King DL, Delfabbro PH, Griffiths MD, Gradisar M. Cognitive-behavioral approaches to outpatient treatment of internet addiction in children and adolescents. J Clin Psychol 2012 Nov;68(11):1185-1195. [doi: 10.1002/jclp.21918] [Medline: 22976240]

15. King DL, Delfabbro PH. The cognitive psychology of Internet gaming disorder. Clin Psychol Rev 2014 Jun;34(4):298-308 [FREE Full text] [doi: 10.1016/j.cpr.2014.03.006] [Medline: 24786896]

16. King DL, Delfabbro PH. Is preoccupation an oversimplification? A call to examine cognitive factors underlying internet gaming disorder. Addiction 2014 Sep 15;109(9):1566-1567. [doi: 10.1111/add.12547] [Medline: 24735375]

17. Wiers RW, Gladwin TE, Hofmann W, Salemink E, Ridderinkhof KR. Cognitive Bias Modification and Cognitive Control Training in Addiction and Related Psychopathology. Clinical Psychological Science 2013 Jan 09;1(2):192-212. [doi: 10.1177/2167702612466547]

18. Jeong EJ, Kim DJ, Lee DM. Why Do Some People Become Addicted to Digital Games More Easily? A Study of Digital Game Addiction from a Psychosocial Health Perspective. International Journal of Human–Computer Interaction 2016 Sep 20;33(3):199-214. [doi: 10.1080/10447318.2016.1232908]

19. James RJE, Tunney RJ. The relationship between gaming disorder and addiction requires a behavioral analysis. J Behav Addict 2017 Sep 01;6(3):306-309 [FREE Full text] [doi: 10.1556/2006.6.2017.045] [Medline: 28816498]

20. King DL, Herd MCE, Delfabbro PH. Tolerance in Internet gaming disorder: A need for increasing gaming time or something else? J Behav Addict 2017 Dec 01;6(4):525-533 [FREE Full text] [doi: 10.1556/2006.6.2017.072] [Medline: 29137493]

21. Bekele E, Bian D, Peterman J, Park S, Sarkar N. Design of a Virtual Reality System for Affect Analysis in Facial Expressions (VR-SAAFE); Application to Schizophrenia. IEEE Trans. Neural Syst. Rehabil. Eng 2017 Jun;25(6):739-749. [doi: 10.1109/tnsre.2016.2591556]

22. Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. The Lancet 2007 Apr 14;369(9569):1302-1313. [doi: 10.1016/S0140-6736(07)60368-7] [Medline: 17434406]

23. Vanheusden K, Mulder CL, van der Ende J, van Lenthe FJ, Mackenbach JP, Verhulst FC. Young adults face major barriers to seeking help from mental health services. Patient Educ Couns 2008 Oct;73(1):97-104. [doi: 10.1016/j.pec.2008.05.006] [Medline: 18584997]

24. Schuemie MJ, van der Straaten P, Krijn M, van der Mast CA. Research on presence in virtual reality: a survey. Cyberpsychol Behav 2001 Apr;4(2):183-201. [doi: 10.1089/109493101300117884] [Medline: 11710246]

25. Park S, Hwang HS. Understanding Online Game Addiction: Connection between Presence and Flow. In: Jacko JA, editor. Human-Computer Interaction. Interacting in Various Application Domains. HCI 2009. Lecture Notes in Computer Science. Heidelberg, Germany: Springer; 2009.

XSL·FO

RenderX

26.   Lidstone JSM, Meins E, Fernyhough C. The roles of private speech and inner speech in planning during middle childhood: evidence from a dual task paradigm. J Exp Child Psychol 2010 Dec;107(4):438-451. [doi: 10.1016/j.jecp.2010.06.002] [Medline: 20633894]

27.   Manfra L, Davis KD, Ducenne L, Winsler A. Preschoolers' motor and verbal self-control strategies during a resistance-to-temptation task. J Genet Psychol 2014;175(3-4):332-345. [doi: 10.1080/00221325.2014.917067] [Medline: 25175682]

28.   Su W, Fang X, Miller JK, Wang Y. Internet-based intervention for the treatment of online addiction for college students in China: a pilot study of the Healthy Online Self-helping Center. Cyberpsychol Behav Soc Netw 2011 Sep;14(9):497-503. [doi: 10.1089/cyber.2010.0167] [Medline: 21294635]

29.   Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]

30.   Dong G, Wang L, Du X, Potenza MN. Gender-related differences in neural responses to gaming cues before and after gaming: implications for gender-specific vulnerabilities to Internet gaming disorder. Soc Cogn Affect Neurosci 2018 Nov 08;13(11):1203-1214 [FREE Full text] [doi: 10.1093/scan/nsy084] [Medline: 30272247]

31.   American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Washington, DC: American Psychiatric Association; 2013.

32.   Sheehan DV, Sheehan KH, Shytle RD, Janavs J, Bannon Y, Rogers JE, et al. Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). J Clin Psychiatry 2010 Mar;71(3):313-326. [doi: 10.4088/JCP.09m05305whi] [Medline: 20331933]

33.   Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34 [FREE Full text] [Medline: 9881538]

34.   Donders J. A short form of the WISC--III for clinical use. Psychological Assessment 1997;9(1):15-20. [doi: 10.1037/1040-3590.9.1.15]

35.   Silverstein AB. Two- and four-subtest short forms of the Wechsler Adult Intelligence Scale-Revised. Journal of Consulting and Clinical Psychology 1982;50(3):415-418. [doi: 10.1037/0022-006x.50.3.415]

36.   Wechsler D. Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). San Antonio, TX: NCS Pearson; 2008:816-827.

37.   Chan PA, Rabinowitz T. A cross-sectional analysis of video games and attention deficit hyperactivity disorder symptoms in adolescents. Ann Gen Psychiatry 2006 Oct 24;5:16 [FREE Full text] [doi: 10.1186/1744-859X-5-16] [Medline: 17059614]

38.   Kim EJ, Namkoong K, Ku T, Kim SJ. The relationship between online game addiction and aggression, self-control and narcissistic personality traits. Eur Psychiatry 2008 Apr;23(3):212-218. [doi: 10.1016/j.eurpsy.2007.10.010] [Medline: 18166402]

39.   Pawlikowski M, Brand M. Excessive Internet gaming and decision making: do excessive World of Warcraft players have problems in decision making under risky conditions? Psychiatry Res 2011 Aug 15;188(3):428-433. [doi: 10.1016/j.psychres.2011.05.017] [Medline: 21641048]

40.   Kim NR, Hwang SS, Choi J, Kim D, Demetrovics Z, Király O, et al. Characteristics and Psychiatric Symptoms of Internet Gaming Disorder among Adults Using Self-Reported DSM-5 Criteria. Psychiatry Investig 2016 Jan;13(1):58-66 [FREE Full text] [doi: 10.4306/pi.2016.13.1.58] [Medline: 26766947]

41.   Bahn GH, Shin M, Cho SC, Hong KE. A preliminary study for the development of the assessment scale for ADHD in adolescents: reliability and validity for Conners-Wells' Adolescent self-report scale-short form. Korean J Child & Adol Psychiatr 2001;12(2):218-224.

42.   Park J, Lee WH, Lee S, Kim S, Bahn GH. Reliability and Validity of Korean Version of the Conners Adult Attention Deficit Hyperactivity Disorder Scale in General Population. J Korean Neuropsychiatr Assoc 2013;52(5):342. [doi: 10.4306/jknpa.2013.52.5.342]

43.   Conners CK. Conners' Rating Scales-revised (CRS-R) CRS-R Technical Manual: CRS09M. Toronto, ON: Multi-Health Systems; 1997.

44.   Chon KK, Choi SC, Yang BC. Integrated adaptation of CES-D in Korea. Korean Journal of Health Psychology 2001;6(1):59-76.

45.   Radloff LS. The CES-D Scale. Applied Psychological Measurement 2016 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]

46.   Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]

47.   Baldwin CK, Caldwell LL. Development of the Free Time Motivation Scale for Adolescents. Journal of Leisure Research 2017 Dec 13;35(2):129-151. [doi: 10.1080/00222216.2003.11949987]

48.   Witmer BG, Jerome CJ, Singer MJ. The Factor Structure of the Presence Questionnaire. Presence: Teleoperators & Virtual Environments 2005 Jun 21;14(3):298-312. [doi: 10.1162/105474605323384654] [Medline: 5637250]

49.   Witmer BG, Singer MJ. Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence 1998 Jun 21;7(3):225-240. [doi: 10.1162/105474698565686] [Medline: 5637250]

50.   Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. The International Journal of Aviation Psychology 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]

51.   Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann 1982;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-x] [Medline: 10259963]

52.   Cox WM, Klinger E, Fadardi JS. The motivational basis of cognitive determinants of addictive behaviors. Addict Behav 2015 May;44:16-22. [doi: 10.1016/j.addbeh.2014.11.019] [Medline: 25499579]

53.   Hilgard J, Engelhardt CR, Bartholow BD. Individual differences in motives, preferences, and pathology in video games: the gaming attitudes, motives, and experiences scales (GAMES). Front Psychol 2013;4:608 [FREE Full text] [doi: 10.3389/fpsyg.2013.00608] [Medline: 24058355]

54.   Park SY, Kim SM, Roh S, Soh M, Lee SH, Kim H, et al. The effects of a virtual reality treatment program for online gaming addiction. Comput Methods Programs Biomed 2016 Jun;129:99-108. [doi: 10.1016/j.cmpb.2016.01.015] [Medline: 26860055]

55.   Zhang Y, Ndasauka Y, Hou J, Chen J, Yang LZ, Wang Y, et al. Cue-induced Behavioral and Neural Changes among Excessive Internet Gamers and Possible Application of Cue Exposure Therapy to Internet Gaming Disorder. Front Psychol 2016 May 09;7:675 [FREE Full text] [doi: 10.3389/fpsyg.2016.00675] [Medline: 27242589]

56.   Shin Y, Kim J, Kim M, Kyeong S, Jung YH, Eom H, et al. Development of an effective virtual environment in eliciting craving in adolescents and young adults with internet gaming disorder. PLoS One 2018 Apr 19;13(4):e0195677 [FREE Full text] [doi: 10.1371/journal.pone.0195677] [Medline: 29672530]

57.   Markland D, Ryan RM, Tobin VJ, Rollnick S. Motivational Interviewing and Self–Determination Theory. Journal of Social and Clinical Psychology 2005 Sep 21;24(6):811-831. [doi: 10.1521/jscp.2005.24.6.811] [Medline: 5637250]

58.   Morina N, Ijntema H, Meyerbröker K, Emmelkamp PMG. Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. Behav Res Ther 2015 Nov;74:18-24. [doi: 10.1016/j.brat.2015.08.010] [Medline: 26355646]

59.   Jinjakam C, Hamamoto K. Simulator sickness in immersive virtual environment. 2013 Presented at: Biomedical Engineering International Conference (BMEiCON); December 5-7, 2012; Ubon Ratchathani, Thailand. [doi: 10.1109/bmeicon.2012.6465465]

60.   Moss JD, Muth ER. Characteristics of head-mounted displays and their effects on simulator sickness. Hum Factors 2011 Jun 27;53(3):308-319. [doi: 10.1177/0018720811405196] [Medline: 21830515]

61.   Sharples S, Cobb S, Moody A, Wilson JR. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. Displays 2008 Mar;29(2):58-69. [doi: 10.1016/j.displa.2007.09.005]

Abbreviations

ADHD: Attention Deficit Hyperactivity Disorder
BAI: Beck Anxiety Inventory
CASS-S: Conners-Wells' Adolescent Self-Report Scale
CBT: cognitive behavioral therapy
CES-D: Center for Epidemiological Studies Depression Scale
CSQ: Client Satisfaction Questionnaire
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th edition
FTMS-A: Free Time Motivation Scale for Adolescents
IAT: Internet Addiction Test
IGD: internet gaming disorder
MET: motivational enhancement therapy
MINI-KID: Mini-International Neuropsychiatric Interview for Children and Adolescents
PQ: presence questionnaire
SSQ: Simulator Sickness Questionnaire
VAS: visual analog scale
VR: virtual reality

Original Paper

A Cognitive-Based Board Game With Augmented Reality for Older Adults: Development and Usability Study

Yen-Fu Chen[1], PhD; Sylvia Janicki[1], MLA, MUP

Department of Media Design, Tatung University, Taipei, Taiwan

Corresponding Author:
Yen-Fu Chen, PhD
Department of Media Design
Tatung University
40, Sec 3, Zhongshan N Rd
Taipei, 104
Taiwan
Phone: 886 221822928 ext 6732
Email: yfchen@gm.ttu.edu.tw

Abstract

Background: Older adults in Taiwan are advised to adopt regular physical and social activities for the maintenance of their cognitive and physical health. Games offer a means of engaging older individuals in these activities. For this study, a collaborative cognitive-based board game, Nostalgic Seekers, was designed and developed with augmented reality technology to support cognitive engagement in older adults.

Objective: A user study of the board game was conducted to understand how the game facilitates communication, problem solving, and emotional response in older players and whether augmented reality is a suitable technology in game design for these players.

Methods: A total of 23 participants aged 50 to 59 years were recruited to play and evaluate the game. In each session, participants' interactions were observed and recorded, then analyzed through Bales' interaction process analysis. Following each session, participants were interviewed to provide feedback on their experience.

Results: The quantitative analysis results showed that the participants engaged in task-oriented communication more frequently than social-emotional communication during the game. In particular, there was a high number of answers relative to questions. The analysis also showed a significant positive correlation between task-oriented acts and the game score. Qualitative analysis indicated that participants found the experience of playing the game enjoyable, nostalgic objects triggered positive emotional responses, and augmented reality technology was widely accepted by participants and provided effective engagement in the game.

Conclusions: Nostalgic Seekers provided cognitive exercise and social engagement to players and demonstrated the positive potential of integrating augmented reality technology into cognitive-based games for older adults. Future game designs could explore strategies for regular and continuous engagement.

Introduction

Background

The growing older population has become a major source of concern in the past decade due to the demographic transition occurring in many developed countries [1]. One of the Sustainable Development Goals (SDGs) listed by the United Nations is to ensure healthy lives and promote well-being for all ages [2]. To support healthy aging and well-being, a variety of factors should be considered, for example, preventing injury or illness, maintaining physical and cognitive function, retaining intimate human relationships, and actively participating in social activities [3]. Studies have shown that the cognitive health of older populations is related to their performance in daily activities and that the presence of cognitive impairments strongly impacts older individuals' level of independence [4-6]. Thus, it is important to assist older adults in maintaining their cognitive

health in order to support their ability to perform daily activities and live independently. Taiwan is one of the fastest-aging societies in the world [1]. The Taiwanese government has undertaken measures to address issues relating to the growing older population. Some public policies that support older persons in Taiwan focus on the development of a society for active aging [7]. These policies suggest that Taiwanese older adults adopt regular physical and social activities for the maintenance of their cognition and health. However, active and continuous engagement in physical and social activities can be challenging for older adults.

Games may offer a means of engaging older persons in regular health-promoting social and physical activities because they can also provide a source of entertainment and engagement. Studies have shown that gaming can produce general cognitive and perceptual improvements, making games a promising approach to preventing or slowing down cognitive decline [8-11]. In recent years, different types of games have been used to investigate the effects of game play on the cognition of older populations [12-14]. These studies have demonstrated that games have the potential to serve as effective cognitive training tools that can benefit cognitive and perceptual abilities in older adults. With growing technological development, digital games could also be easily incorporated as a daily activity for older players. Recent reports from the United States and the United Kingdom have shown that the number of older players playing digital games is on the rise [15,16].

As older populations continue to grow, older players could potentially become a very large customer base in the gaming market. However, the older population has not typically been targeted as an audience in the design of digital games [17]. Currently, digital games are mainly designed for adolescents and young adults and do not consider the physical limitations or psychological preferences of older people. As a result, many digital games are unenjoyable or unsuitable for older adults because design considerations ignore the needs of older populations. For example, the size of the objects on the screen may be too small or the movements and reactions required of the user may be too rapid [18]. Nevertheless, there have been some studies that have investigated digital game preferences of older adults. For example, older adults dislike violence in digital games [15] and prefer intellectually stimulating games, such as puzzle games [19]. Games and game system designs that ignore the needs of older gamers may exclude a large number of potential players [20]. Therefore, it is important for game designers and developers to understand the capabilities, limitations, and interests of older gamers.

In addition to understanding older adults' limitations and preferences in games, an equally important aspect of developing digital games for them is understanding their acceptance of new technology in games and their willingness to learn a new form of technology to play games. The technology of mobile devices (such as smartphones and tablets) is well established, commonly used, and affordable for the general public. Mobile devices seem to be readily adopted by older adults in their daily lives [21]. Digital technology, such as augmented reality (AR), has been widely applied to mobile devices, and its applications are on the rise [22]. AR integrates virtual information with information

in the physical environment [23]. Like digital games, AR is usually associated with younger user groups, and older adults are often excluded in the development of AR technology. However, AR applications have the potential to improve older adults' social relationships, overall well-being, and quality of life and to support their independence [23,24]. Studies show that even though interacting with digital technology may present unique challenges to older users [25], employing new technology in games can enhance older players' immersion and flow [26]. As such, mobile devices with AR technology used for gaming should offer older adults a new and enhanced gaming experience. However, there have been limited studies examining the effects of AR technology in gaming for older players.

For this study, we designed, built, and evaluated a cognitive-based cooperative board game that integrates AR technology and targets Taiwanese older players. We examined how the game facilitates communication and evokes emotional responses and how players responded to the use of new technology in the game. We studied Taiwanese individuals aged 50 to 59 years, as they represent the next generation of older adults in Taiwan and because of their familiarity with mobile devices in daily life [21]. The content of the AR game was developed to reflect Taiwanese culture and history.

Literature Review

Cognitive Health Games

Cognitive performance in working memory, executive function, semantic memory, and logical reasoning could be enhanced by playing games [27,28]. Serious games are games designed to achieve purposes beyond entertainment [29]. They are designed to be used as an educational tool to inform or as a training tool to encourage behavioral changes. Cognitive health games are a form of serious games that support cognitive function by integrating tasks that may help to prevent or slow down cognitive decline and providing players with information relevant to maintaining cognitive health [8,9]. In general, cognitive health games have three main purposes: (1) a preventative purpose intended for healthy players to prevent cognitive decline, (2) a therapeutic purpose intended for cognitively impaired players to maintain mental activity and slow down cognitive decline, and (3) an informational purpose intended to provide the player with formal or informal cognitive assessments and educate the player about different cognitive disorders [30]. Cognitive health games include cognitive training games, which aim to improve a player's cognitive abilities through cognitive tasks, and cognitive screening games, which are used to assess the player's performance in cognitive tasks [31].

Though digital games are becoming more prolific, studies seem to suggest that board games can attain emotional satisfaction more effectively. Using Norman's 3 emotional design levels—the visceral level, behavioral level, and reflective level—as a measure [32], a study comparing video games to board games found that the satisfaction levels declined on all 3 levels for study participants when playing digital games [33]. Another study found that older populations specifically are less drawn to video games than young adults and are more inclined to play board games [26]. The physical appearance and

tangibility of board games may make them more legible and memorable to players and reduce errors in operation. Players can experience intimacy, vivid imagery, sympathetic responses, and social satisfaction while playing board games with others in person [33]. Because of the face-to-face interaction during play, board games give players more social connection and social exchange than virtual or digital games. This may explain why playing board games not only decreases cognitive decline in older adults but reduces the rate of depression as well [34]. Because of these advantages, board games appear to be an appropriate means to support cognitive health in older populations. As such, our study focuses on the development of a cognitive-based board game for older adults for cognitive training.

Nostalgia as the Cognitive Element of Older Adults' Gameplay

Cognitive responses are commonly shown to be affected by feelings of nostalgia [35,36]. Nostalgia refers to a general liking or favorable affect toward objects or events of the past [36]. It differs from memory in that memory refers to the ability to remember information, experiences, and people, whereas nostalgia consists of remembering and reflecting [37], which includes emotional reactions that can be positive (ie, warmth, tenderness, joy, elation) or negative (ie, loss, fear, sadness) [38].

Although nostalgia can be triggered by negative emotions such as loneliness and depression [37], it typically serves as a repository of positive affect [39] and has also been characterized as a joyous experience that provide "a feeling of elation" [40]. Nostalgia may provide a positive view of the past that can help increase players' feelings of social connectedness and enhance meaning in their lives [41]. For populations especially vulnerable to social isolation, such as first-year boarding students, immigrants, and older adults, experiencing feelings of nostalgia may help overcome feelings of loneliness [42]. In general, there are two different types of nostalgic reactions: personal and historical nostalgia [43]. Personal nostalgia refers to reactions generated from a memory of one's past ("the way I was"), whereas historical nostalgic reactions are related to a general period of time in history that could include a time before one was born (also known as historical or communal nostalgia, or "the way it was") [43,44]. The occurrence of nostalgia appears to vary with age: nostalgia levels tend to be higher among young adults, then drop in middle age and rise again during old age [45]. In general, nostalgia is set off by two types of triggers. One type occurs internally and usually arises from feelings of psychological distress, such as loneliness [39], meaninglessness [46], and boredom [47]. When people have negative feelings, they may turn their thoughts toward the past to deal with their current discomfort. Another type of trigger occurs externally through sensory stimuli in one's environment that reminds an individual of their past, such as smells [48], tastes [49], the sight of different objects (often childhood related) [50], and music [51].

For this study, we incorporated imagery of nostalgic objects into the visual materials of the game to evoke memories of players' childhoods that might lead to positive emotions and stimulate cognitive response in participants. Nostalgic objects were selected to be external visual triggers to elicit both a historic nostalgic reaction, in which participants are reminded of a time period in the past that they lived through, and a personal nostalgic reaction, in which the objects might also remind participants of their own stories from childhood. The details of the selection of nostalgic objects for the game is discussed in a later section.

AR Technology for Older Adults

Although AR has existed for about three decades, AR systems have only become affordable and available to the general public because of the recent proliferation of mobile devices [52]. AR is defined as a system or visualization technique that fulfils three main criteria: (1) a combination of real and virtual worlds, (2) real-time interaction, and (3) accurate three-dimensional (3D) registration of virtual and real objects [53]. AR differs from virtual reality (VR) in that VR allows the user to be completely immersed inside an artificial environment, whereas AR is a real-time technology, whereby a physical environment is augmented by adding or embedding virtual information [54]. It allows for a form of information presentation that can more efficiently convey abstract concepts and in turn help improve the spatial and cognitive abilities of its users [55].

Most applications with AR are available using mobile devices that have built-in cameras, GPS sensors, and internet access to embed real-world environments with dynamic, context-aware, and interactive digital content. Currently, 2 forms of AR are available to app developers: vision-based and location-aware AR [56]. Vision-based AR shows the digital content on the screen of mobile devices when users point the built-in camera at an object, such as a two-dimensional (2D) graphic or quick response code. Location-aware AR shows the digital content on the screen when users move through a physical area with the GPS-enabled mobile device. Many studies in the areas of gerontology and human-computer interaction have focused on the use of AR technology to support aging-related needs. For example, Korn et al [57] developed a system combining AR and gamification to support older adults' rehabilitation activities. Kim and Dey [58] developed an AR navigation display system on vehicle windshields to assist older drivers, showing a significant reduction in the number of navigational errors made by users. These studies suggest that AR technology can be a supportive tool to respond to age-related changes in cognitive and physical abilities when designing for older populations. In recent years, AR technology has been applied to the design of various board games to enhance players' gaming experiences [59,60]. Boletsis and McCallum [61] developed the cognitive augmented reality cubes system, which is a collection of cognitive-based mini-games combining AR and physical objects that is designed to prevent cognitive decline in older players through frequent play. Their research has shown that integrating AR technology into games can enhance engagement for older adults [61].

In this study, we integrated vision-based AR technology into the design of a cognitive-based collaborative board game to explore how older adults react to the use of AR technology in games and examine whether AR technology can enhance their gaming experience.

XSL·FO

RenderX

The Design of a Cognitive-Based Board Game With AR

In response to the rapidly growing older population in Taiwan and the rising need to promote active aging and cognitive health in older populations, we designed Nostalgic Seekers, a cognitive-based collaborative board game that integrates AR technology and targets Taiwanese players aged 50 and older. The game consists of physical puzzle pieces designed to be assembled on a table and vision-based AR technology that responds to combinations of puzzle pieces that form specific 2D graphics. An AR app was developed for the game so players could point the built-in camera of their mobile device at a 2D graphic of the board game to conjure a digital 3D model. These 3D AR models take the form of various objects from a period that the targeted players might associate with their childhood. The game is designed to be a cognitive training tool for older adults. Following its design and development, we conducted a user test to examine the cognitive effects of the game, presented in the "Methods" section. With the development of this game, we wanted to examine if a board game integrating AR technology was an appropriate tool to engage older Taiwanese adults and how our game supports the cognitive health of this population.

The Selection of Nostalgic Objects for the Board Game

Nostalgic Seekers was designed for Taiwanese individuals aged 50 to 59, considered the next generation of Taiwanese older adults. To elicit feelings of nostalgia in these players, the AR-projected artifacts in the game resemble various products from the 1960s, 1970s, and 1980s, covering the period that represents the childhood of our target players. These 3 decades also represent an era that brought about rapid industrialization and economic growth in Taiwan. During this time, many national development projects were proposed and undertaken. With prolific government planning efforts and the rise of universal education, Taiwan made huge advances in industry and agriculture and raised living standards substantially [62]. Taiwanese people began to experience the convenience of industrialization as families started to buy and use diverse industrialized products, including home electronic products and private vehicles. This era marked a rapid change in the lifestyles of Taiwanese people, representing a critical period in Taiwanese modern history that is referred to as the 'Taiwan Economic Miracle' [62]. As a result, many products from the 1960s, 1970s, and 1980s remain memorable and iconic of a unique time in history to Taiwanese individuals older than 50 years. We selected a range of products from this era to represent the nostalgic objects in Nostalgic Seekers. Two approaches were used to determine which products were most suitable to incorporate into the game: (1) a Likert-scale questionnaire showing images of a wide range of products from the 1960s, 1970s, and 1980s, as seen in Multimedia Appendices 1 and 2, and (2) a brief individual interview following the questionnaire.

Through snowball sampling, a total of 30 participants aged 50 to 59 (10 men and 20 women) were recruited to take part in the initial questionnaire and interviews for the selection of nostalgic objects. The questionnaire was conducted electronically, and images were shown to participants on a screen. In the questionnaire, 93 images of products from the 1960s, 1970s, and 1980s were presented, including home electronic products, vehicles, toys, furniture, accessories, and daily use products. Participants were asked to observe the image, then rate their familiarity with the object on a Likert scale from 1 to 7 (1=strongly disagree; 2=disagree; 3=slightly disagree; 4=neutral; 5=slightly agree; 6=agree; 7=strongly agree). Each participant was given 15 minutes to fill out the questionnaire. After completing the questionnaire, a brief interview was conducted with each participant in which participants provided their thoughts on the objects. Each individual interview lasted 25 to 30 minutes. Interviews were recorded with a voice recorder with participants' consent.

Average scores for each object were calculated following the questionnaire. Scores ranged from 4.50 to 6.63, with a mean score of 5.83. Moreover, the interviews revealed the different emotional reactions participants had to the various objects. Images of toys reminded participants of sentimental memories from childhood, bringing some participants back to the good times they shared with their childhood friends. For example, one male participant said:

> Seeing these toys gives me the feeling of joy. They remind me of the fun times I had with my old friends. I was so happy.

Another participant said:

> When I see these toys, I feel like I am back in my childhood and playing with other children.

Images of toys reminded participants of carefree memories of their childhood playing with others and brought a positive feeling of nostalgia. In addition, images of daily use products, such as toiletries, kitchenware, and cleaning products, though seemingly mundane, reminded participants of the details in their everyday lives, bringing them back to a time in their past. As one participant stated:

> I still remember the smell of Ming Sing Floral Water because my family always used it in the morning. It reminds me of the wonderful time I had with my mom and sisters.

Another participant said:

> These daily use products were seen in every family 30 years ago. They were cheap but useful.
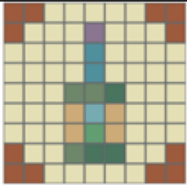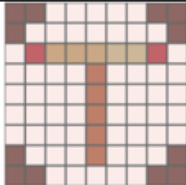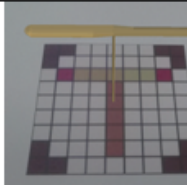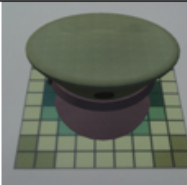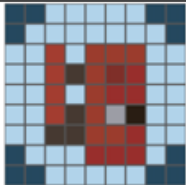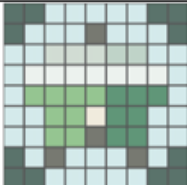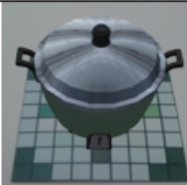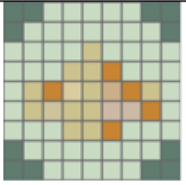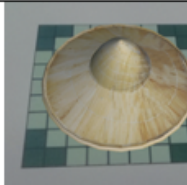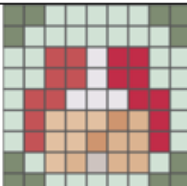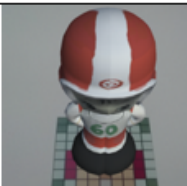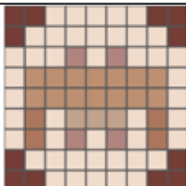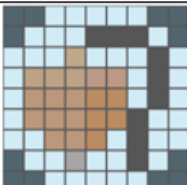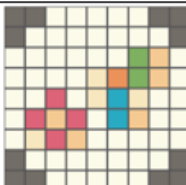
These daily necessities represent different facets of people's lives during the 1960s, 1970s, and 1980s, eliciting nostalgic feelings in participants. Other objects, such as home electronic products, vehicles, furniture, and accessories, also evoked memories in participants. However, they had to be especially iconic and representative of the time for participants to have specific associations with them and consider them nostalgic objects.

Based on this investigation, the following 10 objects were selected for the board game: a traditional Taiwanese perfume, Ming Sing Floral Water (average score of 6.63); a peaked cap (average score of 6.56), an automatic cooker and steamer (average score of 6.56), a toy figurine of Tatung Boy (average score of 6.50); a spinning top (average score of 6.46); a

bamboo-copter (average score of 6.43); a rotary phone (average score of 6.40); an Asian conical hat (average score of 6.40); school desks and chairs (average score of 6.40); and a Japanese candy, Konpetio (average score of 6.40). The 10 selected objects were modeled in Autodesk Maya (Autodesk Inc), a 3D modeling software, and incorporated into the game as the AR images. Using the shape, form, and color of each nostalgic object as guidelines, a 2D pixel art graphic was created to represent each of the 10 objects in an abstracted form. Collectively, these 2D graphics became the puzzle pieces in the game. Each completed 2D graphic represents a 2D target for the vision-based AR in the game. Figure 1 shows the graphics of the 2D pixel art and the 3D models of the 10 objects.

Figure 1. The graphics in pixel art and the 3D models of the 10 objects in Nostalgic Seekers. 2D: two-dimensional; 3D: three-dimensional.

The Rules of the Board Game

Nostalgic Seekers was designed to be a collaborative puzzle game for 4 to 5 players. The components of the game include 5 player tokens, 1 base tile, and 60 puzzle tiles. As shown in Figure 2, these 60 puzzle tiles come in 4 different shapes: the small square tile, the rectangle tile, the L-shaped tile, and the large square tile. Each puzzle tile includes part of the 2D pixel graphic for one or more of the nostalgic objects, so the players have to put the puzzle tiles together to construct the complete 2D pixel image of a nostalgic object (Figure 3).

Figure 2. The 4 different types of puzzle tiles and the pixel graphics on the tiles.

Figure 3. The red square area shows a complete pattern that was constructed using different types of puzzle tiles.

Each player has one player token representing a treasure seeker looking for nostalgic artifacts in the game. The player tokens are placed on top of the base tile, which sits at the center of the play area. The puzzle tiles are placed in piles at the edge of the play area away from the base tile. Puzzle tiles of the same shape should be put in respective piles. For each turn, a player can perform 2 of the following 4 actions: (1) move the player token by 1 position, (2) pick 1 puzzle tile from one of the 4 piles and put the puzzle tile next to the tile that has the player token, (3) move a puzzle tile in the play area to complete the pattern, and (4) inspect a completed pattern of the puzzle tiles via a mobile device, such as a mobile phone or tablet, using the associated app.

After a player has taken 2 actions in their turn, the next player goes. During the game, players have to put the tiles together, then inspect the combined tiles via the AR technology app with a mobile device, such as a mobile phone or tablet. If the tiles are assembled correctly and form the accurate 2D pixel art, a 3D object will appear on the screen along with the game score (Figure 4). Players can discuss strategies and moves with each other during the game, but only 1 player can take an action at one time. Players have 20 minutes to complete the game.

Figure 4. The three-dimensional object of one of the nostalgic artifacts, an automatic cooker and steamer, is presented on the screen of a mobile device when the pattern of the puzzle tiles is assembled correctly.

Methods

To evaluate Nostalgic Seekers, 23 of the original participants who gave feedback on the nostalgic objects volunteered to participate. A total of 13 women and 10 men aged 50 to 59 years participated in the evaluation workshops. In consideration of participants' variable availability, 7 game-testing workshops were arranged on different dates, and each workshop had 4 participants (Figure 5). To keep the number of participants in each workshop the same, 5 participants were invited to participate in 2 different workshops. However, they did not play with any other participant twice. Each workshop lasted 1 hour, including 30 minutes for playing the game. This study was approved by the research ethics committee. We obtained written permission to use images of individuals included in this paper.

At the beginning of the game-testing workshops, the research team explained the rules of Nostalgic Seekers in detail to the participants. Two digital cameras and one sound recorder were set up in each workshop to record the participants' interactions during the game. After the game ended, each participant gave their feedback on the game through structured interviews, which were recorded with a voice recorder.

Observing players' interactions with each other in the game play is a useful approach to evaluating the game design because players engage in both verbal and nonverbal communication. Bales' interaction process analysis (IPA) is a method for the analysis of communication processes among small groups [63]. IPA consists of 12 complementary paired process categories for communication acts [63,64]. These 12 process categories are further subdivided into 4 major functions: social-emotional acts for positive reaction, social-emotional acts for negative reaction, task-oriented acts for questions, and task-oriented acts for attempted answers. Table 1 shows the system of process categories in the IPA.

Figure 5. Four participants played the game in one of the game-testing workshops.

Table 1. System of process categories in the interaction process analysis.

Function and process categories	Paired processes and central problems addressed
Social-emotional: positive reaction	
1. Shows solidarity, raises other's status, gives help, rewards	1 & 12; integration
2. Shows tension release, jokes, laughs, shows satisfaction	2 & 11; tension management
3. Agrees, shows passive acceptance, understands, concurs, complies	3 & 10; decision
Task area: attempted answers	
4. Gives suggestion or direction, implying autonomy for other	4 & 9; control
5. Gives opinion, evaluation, or analysis, expresses feeling or wish	5 & 8; evaluation
6. Gives orientation or information, repeats, clarifies, confirms	6 & 7; orientation
Task area: questions	
7. Asks for orientation, information repetition, or confirmation	7 & 6; orientation
8. Asks for opinion, evaluation, analysis, or expression of feeling	8 & 5; evaluation
9. Asks for suggestion, direction, or possible action	9 & 4; control
Social-emotional: negative reaction	
10. Disagrees, shows passive rejection or formality, withholds help	10 & 3; decision
11. Shows tension, asks for help, withdraws out of field	11 & 2; tension management
12. Shows antagonism, deflates other's status or defends or asserts self	12 & 1; integration

A total of 6 of the 12 process categories are related to social-emotional acts, with 3 positive and 3 negative types of expressions representing sociability and affect. Positive social-emotional content expresses solidarity or friendliness (category 1), tension relief or dramatizing (category 2), or agreement and understanding (category 3). Negative social-emotional content shows disagreement and passive rejection (category 10), tension (category 11), and antagonism (category 12). Categories 1 and 12 are related to issues of integration. Categories 2 and 11 are related to issues of tension

management. Categories 3 and 10 are related to issues of decision making.

The other 6 process categories are task-oriented acts, including the exchange of questions and responses to complete a task. The 3 types of questions in task-oriented acts are asking for task information or orientation (category 7), asking for an opinion (category 8), and asking for a suggestion (category 9). The 3 types of answers in task-oriented acts are giving a suggestion or command (category 4), giving an opinion (category 5), and giving task information or orientation (category 6). Categories

4 and 9 relate to problems of control. Categories 5 and 8 relate to problems of evaluation. Categories 6 and 7 relate to problems of orientation. The IPA uses a single act as the unit for coding and analysis. An act is a single sentence or its equivalent, such as nonverbal communication or an exchange that may be understood by other members.

Sometimes participants' verbal communication was also accompanied by nonverbal behavior. To avoid coding the same acts twice, when both verbal and nonverbal communication occurred, the coding only counted verbal communication. Two coders were trained in accordance with the IPA coding scheme. Cohen κ was used to assess the interrater reliability for the agreement between the 2 coders. The Cohen κ values were greater than 7 for all 7 workshops and thus satisfactory in all instances. Specifically, the kappa index (κ) for each workshop, in order of the workshop, was 0.76, 0.86, 0.81, 0.88, 0.83, 0.79, and 0.89.

Structured interviews were conducted with each participant following each game and recorded with a voice recorder. Participants were asked to answer questions related to their experience of the puzzle game, specifically how they felt about the design and goals of the board game, the process of collaborative problem solving, the use of AR technology, and their reactions to the nostalgic objects.

Results

Participants' Communication Acts

After the coding was complete, a total of 2021 communication acts were recorded from these 7 workshops; 278 total communication acts were observed in the first workshop, 144 in the second workshop, 213 in the third workshop, 132 in the fourth workshop, 349 in the fifth workshop, 465 in the sixth workshop, and 431 in the seventh workshop. Table 2 shows the observed frequencies of participants' communication acts in these 7 game-testing workshops.

Table 2. The observed frequencies of participants' communication acts in the game-testing workshops.

Game test workshop	Social-emotional acts		Task-oriented acts		Game score
	Positive	Negative	Attempted answers	Questions	
1st	20	27	165	66	2
2nd	12	15	78	39	0
3rd	10	2	139	62	2
4th	27	4	75	26	0
5th	42	4	208	95	2
6th	33	9	293	130	3
7th	60	10	254	107	3

Participants' social-emotional acts and task-oriented acts were analyzed with a chi-square goodness-of-fit test to assess the interpersonal interaction during the game test. The value (χ^2_3=7.81; α=.05) on the chi-square distribution table indicated that when a chi-square value was greater than 7.81, there was a statistically significant difference in the data. The chi-square tests testing the observed frequencies of participants' social-emotional acts and task-oriented acts achieved statistical significance in all workshops (first workshop: χ^2_3=162.65 > 7.81; second workshop: χ^2_3=77.5 > 7.81; third workshop: χ^2_3=223.98 > 7.81; fourth workshop: χ^2_3=81.52 > 7.81; fifth workshop: χ^2_3=270.7 > 7.81; sixth workshop: χ^2_3=428.93 > 7.81; seventh workshop: χ^2_3=286.6 > 7.81).

These results indicate that participants performed task-oriented acts much more frequently than social-emotional acts when playing the game. For the frequency of participants' task-oriented acts, the frequency of attempted answers was more than 2 times greater than the frequency of questions in each workshop.

The analysis results shown in Table 3 indicate that the frequency of task-oriented acts, namely the communication acts of questions (R^2=0.834; F_1=25.184) and attempted answers (R^2=0.870; F_1=33.535), significantly predicted the game score, with greater frequencies of each correlating with higher scores. However, social-emotional acts of positive reactions (R^2=0.303; F_1=2.172) and negative reactions (R^2=0.000; F_1=0.002) did not correlate to the game score. Communication acts of questions accounted for 83% and attempted answers accounted for 87% of the total variance in the game score, showing that there was a significant relationship between the game score and participants' task-oriented acts. This suggests that the game encourages participant discussion and cooperation for problem solving, providing participants a chance to practice their cognitive skills. In particular, the high number of answers relative to questions signals that participants were collaborating to find solutions together. Specifically, in the game-testing workshops, we saw participants not only communicating to solve the puzzle together but also collaborating to figure out the use of the AR technology.

Table 3. Simple linear regression analysis results of participants' communication acts related to the game scores.

Dependent variable and independent variables	R	R^2	β	t test (df)	F test (df)	P value
Game score[a]						
Positive reactions	0.550	0.303	0.039	1.474 (1)	2.172 (1)	.20
Negative reactions	0.020	0.000	0.003	0.044 (1)	0.002 (1)	.97
Questions	0.913	0.834	0.031	5.018 (1)	25.184 (1)	.004
Attempted answers	0.932	0.870	0.012	5.491 (1)	33.535 (1)	.002

[a]Dependent variable.

On the other hand, for participants' social-emotional acts, the relative frequencies of negative acts and positive acts varied. In the first and second game-testing workshops, the number of negative acts exceeded positive ones. However, in the remaining workshops, the frequency of positive acts was substantially higher than the frequency of negative acts. In total, across all workshops, the number of positive acts accounted for almost triple the number negative acts. The low number of social-emotional acts, particularly negative ones, signals that the collaborative problem-solving process in the game, including learning a new technology, was largely a rational process and did not trigger many negative emotions or frustrations in participants.

Participants' Feedback on the Game

Following the game in each workshop, participants were asked to provide feedback on the process and experience of playing Nostalgic Seekers through structured interviews. The participant interviews were recorded with a voice recorder and transcribed, then analyzed with deductive analysis to examine the participants' reactions to the puzzle game and collaborative process, their acceptance of AR technology, and their reactions to the theme of nostalgia.

Overall, participants expressed positive experiences with the game. They enjoyed playing the game (16/23, 70%), found the game design intuitive (20/23, 87%), and felt positively about the collaborative process (16/23, 70%). Over half of the participants were especially intrigued by the use of AR technology (12/23, 52%). One female participant said:

> The goal of this game is easy to understand. It's a collaborative puzzle game for us, and we need to discuss. I think the experience of gameplay is great because it encourages us to talk. It also reminds me that we are a team… I was surprised by the use of AR technology in the game. It is stunning because AR is the main part of this game. We may be aware of something on the tiles, but we still have no idea what it is without AR.

This shows that the discussions and communication that occurred in the game through the collaborative game-playing process can have a positive social impact and that the AR component added an element of excitement and engagement that was attractive to participants. Another male participant stated:

> For me, AR technology is a new technology that I am interested in… Even though this game is a board

> game, it cannot be played without AR. This game gave me a good experience with AR and makes me want to try other applications on mobile devices with AR technology!

Participants' feedback revealed that, overall, they not only were willing to learn AR and accepting of the technology but also seemed excited and fascinated by the AR component in the game.

Additionally, participants also described some of the challenges in playing Nostalgic Seekers that were overcome through the collaborative process. In the game, participants had to put puzzle tiles together to construct different graphics of nostalgic objects in pixel art form before they could inspect the graphics via the AR technology with their mobile devices. Although participants were given reference graphics for the different images they were asked to construct, identifying the graphics on the puzzle tiles and assembling them was challenging for some participants. One female participant shared her experience with the pixel art graphics:

> These graphics were not easy to identify, even though I knew these graphics were the abstracted form of some artifacts that I should be familiar with. We spent a lot of time putting the right tiles together… It was a big challenge in the game for us.

The level of difficulty in assembling the puzzle required participants to discuss, collaborate, and use trial and error to problem solve, which was the intention of the game. The same participant continued to say:

> But when we inspected the graphics via mobile phone and got a 3D object on the screen and scored one point, I can't describe the feeling…a sense of achievement! I am so proud of my team.

This participant reveals that even though constructing the puzzle was challenging and time-consuming, the difficulty made the accomplishment more rewarding. Another female participant described a similar experience:

> It seems the graphics in pixel art should not be difficult to recognize, but it was difficult to put these tiles together correctly. The color, the shape… both were challenging for us. However, I got the sense of achievement when we got the 3D object on the screen… it was worth the effort.

Putting the pixel graphics together correctly with the puzzle tiles is central to the game mechanics, and it is the main

XSL·FO

RenderX

cognitive task of the game that serves as cognitive exercise. The challenge in the game not only provides cognitive training but also gives participants a feeling of reward and sense of achievement when they eventually solve the problem and score points. In addition, some participants suggested that more specific directions should be given to assist players in playing the game.

Regarding the use of the 3D nostalgic objects with AR technology, participants agreed that these 3D nostalgic objects were a novel and appropriate reflection of the theme of the game. One male participant said:

> *I am familiar with these 3D objects on the screen, and they were a part of my life…. no wonder the name of this game is Nostalgic Seekers!*

Another female participant gave her feedback:

> *Using AR technology to present these objects is interesting. Yes, we are the seekers and trying to find these nostalgic objects. I like the idea of the game, but a fly in the ointment was that we didn't have the chance to talk about the objects.*

This participant suggested that a facilitated or intentional discussion of the 3D objects could be a positive component to add in the game.

Discussion

Summary of Results

Through the design, testing, and both the quantitative and qualitative analyses of the game Nostalgic Seekers, this study examined how 50- to 59-year-old Taiwanese adults interacted and communicated with each other while playing a collaborative board game with AR technology. Our game was designed to be a cognitive training game that provides cognitive tasks [31], which are achieved through solving a puzzle and using a digital app with AR technology. Overall, we saw that participants' communication during the game heavily focused on task-oriented exchanges rather than social-emotional ones, and the number of task-oriented acts positively correlated with higher scores. This finding indicates that the collaborative process of the game was able to facilitate and reward problem-solving communication; thus, it successfully engages players in cognitive exercise. Specifically, we observed participants not only problem solving for the puzzle itself but also troubleshooting the use of AR technology with each other. We also saw that for social-emotional communication, the overall number of positive acts significantly outnumbered negative ones. However, the number of negative acts exceeded positive ones in workshops 1 and 2. This may in part be due to the "alpha player problem," in which a dominant player takes control of collective decision making and creates a negative atmosphere in the game [65]. In addition, we saw that nostalgic objects brought back specific memories and positive emotions in participants, which they were eager to share and discuss with each other, and helped strengthen the interest and engagement of the players during game play. Finally, we saw that the inclusion of AR objects in the game was engaging and exciting for the players rather than intimidating.

Implications

The overall low number of negative social-emotional acts coupled with qualitative findings that showed that players enjoyed the collaborative problem-solving process and were excited by the AR app in the game indicate that the users embraced the use of AR technology in the game and that the learning of it became a part of the collaborative process. As such, we found that introducing a new technology in games, in particular in a collaborative game, can encourage supportive peer learning of the new technology in a fun way and is a cognitively engaging task in itself. Subsequently, there is strong potential in integrating applications of new technologies in board games as a tool to support cognitive health and active aging. There are also implications of using interactive and entertaining processes, such as collaborative games, as a means to teach new technology to older adults and help them become more comfortable with the use of digital technologies.

Consequently, our game shows potential for being used as a tool to support the Taiwanese government's active aging agenda [7], as well as the United Nation's SDGs for healthy living and well-being for all ages [2]. Consistent with other findings we previously referred to in the "Literature Review" section, our study also demonstrated that new technology can in fact enhance "immersion and flow" for older adults during game play [26], despite also presenting some challenges [25]. In fact, we also found that challenges that arose in navigating the technology became a part of the collaborative problem-solving process in the cognitive-based game. Our findings reiterate the potential and importance of digital games marketed to older adults [15,16], which are currently lacking [17], and could serve as a step to filling this gap.

Limitations and Future Work

Some factors related to our study sample, such as participants' experiences playing board games or their familiarity with AR and other digital applications, could have played a role in the gameplay and affected player interactions. However, these characteristics were not distinguished in our study sample. Future studies could explore if the game has different impacts on older adults with variable familiarity with technology or variable experiences with board games, for example.

In addition, a major concern with the active aging agenda is related to continuous engagement. Some critical considerations we have yet to address include how to motivate older individuals to engage in this form of game play regularly to achieve the desired cognitive benefits and how to design games that can be engaging after repeated use.

Some additional questions arose during the study that could be further examined in future studies. For example, how might communication style, interaction, and association to different nostalgic objects in the game differ between male and female players or players from different geographic regions in Taiwan? Moreover, how might acts of communication and interaction vary in a different cultural setting (such as in a Western culture)? Additional studies could also explore how task-oriented acts and social-emotional acts in games affect cognitive function in older adults differently and how the two types of acts may

complement each other in supporting cognitive health in older adults. These questions have strong implications on modifications to the game design to maximize support for cognitive health in older populations. Future studies could also explore how integrating different virtual components, such as AR, into a physical board game might support cognitive health in unique ways.

Feedback from participants provided valuable suggestions on ways to improve Nostalgic Seekers for greater clarity and enjoyment. For example, some participants expressed that the 2D graphics were too difficult to recognize. Incorporating multiple levels of difficulty that can accommodate players with different cognitive abilities would provide greater accessibility and cognitive benefits to all players. Participants also expressed their interest in having more discussions about the nostalgic objects that they found. Incorporating a component of storytelling or facilitated conversation about the nostalgic objects that players find can encourage participants to engage in active remembering and enhance social and emotional interaction and communication, which are currently less prominent in the game. Finally, in its current design, Nostalgic Seekers does not necessarily appeal to repeated use, as the novelty of discovering the nostalgic objects would wear off once players are aware of the objects. Future iterations could consider ways to modify the game so that it remains exciting and entertaining for continuous use.

Conclusions

Games offer a means of engaging older adults in social and physical activities for the maintenance of their cognition and health. As the number of older adults increases, active aging is becoming a growing priority for community well-being. Older players are also becoming a growing customer base in the game market. Mobile devices with AR technology used for gaming can offer a new gaming experience for these players and enhance their immersion in the gameplay. In this study, we designed and tested Nostalgic Seekers, a cognitive-based collaborative board game integrating AR technology for Taiwanese players 50 years and older. This game provides players a chance to use cognitive skills because it requires discussion to problem solve and achieve tasks. The study found that the use of AR technology was enticing to participants; assembling the puzzles was challenging, but participants felt a sense of achievement when they completed a puzzle and were able to see their prize in the AR model.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Original survey form in Chinese.
[PDF File (Adobe PDF File), 6625 KB - games_v8i4e22007_app1.pdf]

Multimedia Appendix 2
Translated survey form.
[PDF File (Adobe PDF File), 7629 KB - games_v8i4e22007_app2.pdf]

References

1. Tsai W. The Growth of Taiwan's Aging Population and Its Socio-Economic Consequences. Am J Chin Stud 2008;15(2):93-105 [FREE Full text]
2. Transforming our world: the 2030 Agenda for Sustainable Development. United Nations General Assembly. 2015. URL: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E [accessed 2020-06-10]
3. Rowe JW, Kahn RL. Successful aging. Gerontologist 1997 Aug;37(4):433-440. [doi: 10.1093/geront/37.4.433] [Medline: 9279031]
4. Hertzog C, Kramer AF, Wilson RS, Lindenberger U. Enrichment Effects on Adult Cognitive Development: Can the Functional Capacity of Older Adults Be Preserved and Enhanced? Psychol Sci Public Interest 2008 Oct;9(1):1-65. [doi: 10.1111/j.1539-6053.2009.01034.x] [Medline: 26162004]
5. Pedone C, Ercolani S, Catani M, Maggio D, Ruggiero C, Quartesan R, GIFA Study Group. Elderly patients with cognitive impairment have a high risk for functional decline during hospitalization: The GIFA Study. J Gerontol A Biol Sci Med Sci 2005 Dec;60(12):1576-1580 [FREE Full text] [doi: 10.1093/gerona/60.12.1576] [Medline: 16424291]

XSL·FO

RenderX

6.   Diehl M, Willis SL, Schaie KW. Everyday problem solving in older adults: observational assessment and cognitive correlates. Psychol Aging 1995 Sep;10(3):478-491 [FREE Full text] [doi: 10.1037//0882-7974.10.3.478] [Medline: 8527068]

7.   Lin Y, Huang C. Aging in Taiwan: Building a Society for Active Aging and Aging in Place. Gerontologist 2016 Apr;56(2):176-183. [doi: 10.1093/geront/gnv107] [Medline: 26589450]

8.   McCallum S, Boletsis C. Dementia Games: A Literature Review of Dementia-Related Serious Games. 2013 Presented at: 4th International Conference on Serious Games Development and Applications (SGDA 2013); Sep 25-27, 2013; Trondheim, Norway. [doi: 10.1007/978-3-642-40790-1]

9.   McCallum S, Boletsis C. A Taxonomy of Serious Games for Dementia. In: Schouten B, Fedtke S, Bekker T, Schijven M, Gekker A, editors. Games for Health. Wiesbaden, Germany: Springer Vieweg; 2013:219-232.

10.   Li R, Polat U, Scalzo F, Bavelier D. Reducing backward masking through action game training. J Vis 2010;10(14):33. [doi: 10.1167/10.14.33] [Medline: 21191129]

11.   Green CS, Bavelier D. Action-video-game experience alters the spatial resolution of vision. Psychol Sci 2007 Jan;18(1):88-94 [FREE Full text] [doi: 10.1111/j.1467-9280.2007.01853.x] [Medline: 17362383]

12.   Basak C, Boot WR, Voss MW, Kramer AF. Can training in a real-time strategy video game attenuate cognitive decline in older adults? Psychol Aging 2008 Dec;23(4):765-777 [FREE Full text] [doi: 10.1037/a0013494] [Medline: 19140648]

13.   Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature 2013 Sep 5;501(7465):97-101 [FREE Full text] [doi: 10.1038/nature12486] [Medline: 24005416]

14.   Belchior P, Marsiske M, Sisco SM, Yam A, Bavelier D, Ball K, et al. Video game training to improve selective visual attention in older adults. Comput Human Behav 2013 Jul 01;29(4):1318-1324 [FREE Full text] [Medline: 24003265]

15.   Nap H, Kort Y, Ijsselsteijn W. Senior gamers: Preferences, motivations, needs. Gerontechnology 2009;8(4):247-262. [doi: 10.4017/gt.2009.08.04.003.00]

16.   Pratchett R. Gamers in the UK: Digital play, digital lifestyles. BBC Creative Research and Development. 2005. URL: https://crystaltips.typepad.com/wonderland/files/bbc_uk_games_research_2005.pdf [accessed 2020-06-10]

17.   Kaufman D, Suave L, Ireland A. Playful Aging: Digital Games for Older Adults. Age Well 4.2 Project. URL: https://agewell-nce.ca/wp-content/uploads/2020/02/AGE-WELL_WP4.2_White-paper_GAMES.pdf [accessed 2020-11-27]

18.   Whitcomb G. Computer games for the elderly. 1990 Presented at: Conference on Computers and the quality of life (CQL '90); Sep 13-16, 1990; Washington, DC.

19.   De Schutter B. Never Too Old to Play: The Appeal of Digital Games to an Older Audience. Games and Culture 2011 Mar;6(2):155-170. [doi: 10.1177/1555412010364978]

20.   McLaughlin A, Gandy M, Allaire J, Whitlock L. Putting Fun into Video Games for Older Adults. Ergonomics in Design 2012 Apr 20;20(2):13-22. [doi: 10.1177/1064804611435654]

21.   Zickuhr K. Generations and their gadgets. Pew Research Center's Internet and American Life Project. 2011. URL: https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2011/PIP_Generations_and_Gadgets.pdf [accessed 2020-11-27]

22.   With 85.2% CAGR, Augmented Reality (AR) Market Size Will Exceed USD: Zion Market Research. Global News Wire. 2018 Jan 05. URL: http://www.globenewswire.com/news-release/2018/01/05/1284018/0/en/With-85-2-CAGR-Augmented-Reality-AR-Market-Size-Will-Exceed-USD-133-78-Bn-by-2021-Zion-Market-Research.html [accessed 2020-11-27]

23.   Liang S. Research Proposal on Reviewing Augmented Reality Applications for Supporting Ageing Population. Procedia Manufacturing 2015;3:219-226. [doi: 10.1016/j.promfg.2015.07.132]

24.   Kurz D, Fedosov A, Diewald S, Guettler J, Geilhof B, Heuberger M. Towards Mobile Augmented Reality for the Elderly. 2014 Presented at: 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); Sep 10-12, 2014; Munich, Germany. [doi: 10.1109/ismar.2014.6948447]

25.   Eggermont S, Vandebosch H, Steyaert S. Towards the desired future of the elderly and ICT: policy recommendations based on a dialogue with senior citizens. Poiesis Prax 2006 May 25;4(3):199-217. [doi: 10.1007/s10202-005-0017-9]

26.   Mahmud A, Mubin O, Shahid S, Martens J. Designing and evaluating the tabletop game experience for senior citizens. 2008 Presented at: 5th Nordic Conference on Human-computer Interaction (NordiCHI '08); Oct 18-22, 2008; Lund, Sweden. [doi: 10.1145/1463160.1463205]

27.   Stern Y, Blumen HM, Rich LW, Richards A, Herzberg G, Gopher D. Space Fortress game training and executive control in older adults: a pilot intervention. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2011 Nov;18(6):653-677 [FREE Full text] [doi: 10.1080/13825585.2011.613450] [Medline: 21988726]

28.   Clarkson-Smith L, Hartley AA. The game of bridge as an exercise in working memory and reasoning. J Gerontol 1990 Nov;45(6):P233-P238. [doi: 10.1093/geronj/45.6.p233] [Medline: 2229947]

29.   Michael D, Chen S. Serious Games: Games that Educate, Train and Inform. 1st ed. Boston, MA: Thomson Course Technology; 2005.

30.   Sawyer B. From cells to cell processors: the integration of health and video games. IEEE Comput Graph Appl 2008 Dec;28(6):83-85. [doi: 10.1109/MCG.2008.114] [Medline: 19004688]

31. Boletsis K. Augmented Reality Serious Gaming for Cognitive Health [dissertation]. Norwegian University of Science and Technology. 2016. URL: http://hdl.handle.net/11250/2398794 [accessed 2020-11-27]

32. Fishwick M. Emotional Design: Why We Love (or Hate) Everyday Things. J American Culture 2004 Jun;27(2):234-234. [doi: 10.1111/j.1537-4726.2004.133_10.x]

33. Fang Y, Chen K, Huang Y. Emotional reactions of different interface formats: Comparing digital and traditional board games. Adv Mech Eng 2016 Mar 30;8(3):168781401664190. [doi: 10.1177/1687814016641902]

34. Dartigues JF, Foubert-Samier A, Le Goff M, Viltard M, Amieva H, Orgogozo JM, et al. Playing board games, cognitive decline and dementia: a French population-based cohort study. BMJ Open 2013 Aug 29;3(8):e002998 [FREE Full text] [doi: 10.1136/bmjopen-2013-002998] [Medline: 23988362]

35. Muehling DD, Sprott DE. The Power Of Reflection: An Empirical Examination of Nostalgia Advertising Effects. J Advertising 2004 Oct;33(3):25-35. [doi: 10.1080/00913367.2004.10639165]

36. Holbrook M, Schindler R. Echoes of the Dear Departed Past: Some Work in Progress on Nostalgia. In: Holman R, Solomon M, editors. Advances in Consumer Research. Vol 18. Provo, UT: Association for Consumer Research; 1991:330-333.

37. Kercher D. Exploring Nostalgia: Does It Have Positive or Negative Effects on Mental Health. Inpathy Bulletin. 2016 Dec 23. URL: https://inpathybulletin.com/exploring-nostalgia-positive-negative-effects-mental-health/ [accessed 2020-06-10]

38. Barrett FS, Grimm KJ, Robins RW, Wildschut T, Sedikides C, Janata P. Music-evoked nostalgia: Affect, memory, and personality. Emotion 2010;10(3):390-403. [doi: 10.1037/a0019006]

39. Wildschut T, Sedikides C, Arndt J, Routledge C. Nostalgia: content, triggers, functions. J Pers Soc Psychol 2006 Nov;91(5):975-993. [doi: 10.1037/0022-3514.91.5.975] [Medline: 17059314]

40. Kaplan HA. The psychopathology of nostalgia. Psychoanal Rev 1987;74(4):465-486. [Medline: 3124149]

41. Cavanagh SR, Glode RJ, Opitz PC. Lost or fond? Effects of nostalgia on sad mood recovery vary by attachment insecurity. Front Psychol 2015;6:773 [FREE Full text] [doi: 10.3389/fpsyg.2015.00773] [Medline: 26113829]

42. More Than Just Being A Sentimental Fool: The Psychology Of Nostalgia. Association for Psychological Science. 2008 Dec 1. URL: https://www.psychologicalscience.org/news/releases/more-than-just-being-a-sentimental-fool-the-psychology-of-nostalgia.html [accessed 2020-06-10]

43. Tierney J. What Is Nostalgia Good For? Quite a Bit, Research Shows. The New York Times. 2013 Jul 8. URL: https://www.nytimes.com/2013/07/09/science/what-is-nostalgia-good-for-quite-a-bit-research-shows.html?pagewanted=all&_r=0 [accessed 2020-06-10]

44. Stern BB. Historical and Personal Nostalgia in Advertising Text: The fin de siècle effect. J Advertising 1992 Dec;21(4):11-22. [doi: 10.1080/00913367.1992.10673382]

45. Lerner RM. Plasticity, person-context relations, and cognitive training in the aged years: A developmental contextual perspective. Dev Psychol 1990;26(6):911-915. [doi: 10.1037/0012-1649.26.6.911.a]

46. Routledge C, Wildschut T, Sedikides C, Juhl J, Arndt J. The power of the past: nostalgia as a meaning-making resource. Memory 2012 Jul;20(5):452-460 [FREE Full text] [doi: 10.1080/09658211.2012.677452] [Medline: 22639901]

47. van Tilburg WAP, Igou ER, Sedikides C. In search of meaningfulness: nostalgia as an antidote to boredom. Emotion 2013 Jun;13(3):450-461 [FREE Full text] [doi: 10.1037/a0030442] [Medline: 23163710]

48. Reid CA, Green JD, Wildschut T, Sedikides C. Scent-evoked nostalgia. Memory 2015;23(2):157-166 [FREE Full text] [doi: 10.1080/09658211.2013.876048] [Medline: 24456210]

49. Supski S. Aunty Sylvie's Sponge: Foodmaking, Cookbooks and Nostalgia. Cult Stud Rev 2013 Feb 11;19(1):28-49. [doi: 10.5130/csr.v19i1.3074]

50. Holbrook MB, Schindler RM. Market segmentation based on age and attitude toward the past: Concepts, methods, and findings concerning nostalgic influences on customer tastes. J Business Res 1996 Sep;37(1):27-39. [doi: 10.1016/0148-2963(96)00023-9]

51. Cheung W, Wildschut T, Sedikides C, Hepper EG, Arndt J, Vingerhoets AJJM. Back to the future: nostalgia increases optimism. Pers Soc Psychol Bull 2013 Nov;39(11):1484-1496 [FREE Full text] [doi: 10.1177/0146167213499187] [Medline: 23928397]

52. Wu H, Lee SW, Chang H, Liang J. Current status, opportunities and challenges of augmented reality in education. Comput Educ 2013 Mar;62:41-49. [doi: 10.1016/j.compedu.2012.10.024]

53. Azuma RT. A Survey of Augmented Reality. Presence Teleoperators Virtual Environ 1997 Aug;6(4):355-385. [doi: 10.1162/pres.1997.6.4.355]

54. Enyedy N, Danish J, Delacruz G, Kumar M. Learning physics through play in an augmented reality environment. Comput Support Learning 2012 Jul 6;7(3):347-378. [doi: 10.1007/s11412-012-9150-3]

55. Laine TH, Nygren E, Dirin A, Suk H. Science Spots AR: a platform for science learning games with augmented reality. Educ Tech Res Dev 2016 Jan 7;64(3):507-531. [doi: 10.1007/s11423-015-9419-0]

56. Dunleavy M, Dede C. Augmented Reality Teaching and Learning. In: Spector J, Merrill M, Elen J, Bishop M, editors. Handbook of Research on Educational Communications and Technology. New York, NY: Springer New York; 2014:735-745.

57. Korn O, Buchweitz L, Rees A, Bieber G, Werner C, Hauer K. Using Augmented Reality and Gamification to Empower Rehabilitation Activities and Elderly Persons. A Study Applying Design Thinking. 2018 Presented at: AHFE 2018:

International Conference on Applied Human Factors and Ergonomics; July 21-25, 2018; Orlando, FL. [doi: 10.1007/978-3-319-94229-2_21]

58. Kim S, Dey A. Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. 2009 Presented at: SIGCHI Conference on Human Factors in Computing Systems; Apr 20-25, 2009; Boston, MA. [doi: 10.1145/1518701.1518724]

59. Andersen T, Kristensen S, Nielsen B, Grønbæk K. Designing an Augmented Reality Board Game With Children: the Battleboard 3D Experience. 2004 Presented at: 2004 Conference on Interaction Design and Children; June 1-3, 2004; Baltimore, MD. [doi: 10.1145/1017833.1017858]

60. Oda O, Lister L, White S, Feiner S. Developing an augmented reality racing game. 2008 Presented at: 2nd International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN '08); Jan 8-10, 2008; Cancun, Mexico. [doi: 10.4108/icst.intetain2008.2472]

61. Boletsis C, McCallum S. Augmented Reality Cubes for Cognitive Gaming: Preliminary Usability and Game Experience Testing. Int J Serious Games 2016 Mar 01;3(1):3-18. [doi: 10.17083/ijsg.v3i1.106]

62. Tsai P. Explaining Taiwan's Economic Miracle: Are the Revisionists Right? Agenda J Policy Anal Reform 1999;6(1):69-82 [FREE Full text]

63. Bales RF. A Set of Categories for the Analysis of Small Group Interaction. Am Sociol Rev 1950 Apr;15(2):257. [doi: 10.2307/2086790]

64. Bales R. Personality and Interpersonal Behavior. New York, NY: Holt, Rinehart & Winston; 1970.

65. Engelstein G, Shalev I. Building Blocks of Tabletop Game Design: An Encyclopedia of Mechanisms. Boca Raton, FL: CRC Press; 2019.

Abbreviations

2D: two-dimensional
3D: three-dimensional
AR: augmented reality
IPA: interaction process analysis
SDG: Sustainable Development Goals
VR: virtual reality

XSL•FO
RenderX

Original Paper

Effects of a Mobile App Called Quittr, Which Utilizes Premium Currency and Games Features, on Improving Engagement With Smoking Cessation Intervention: Pilot Randomized Controlled Trial

Ivan Bindoff[1], BComp (Hons), PhD; Tristan R Ling[1], BComp (Hons), PhD; Peter Gee[1], BPharm; Benjamin Geelan[1], BIS (Hons), PhD; Stuart G Ferguson[2], BPsych, PhD; Gregory M Peterson[1], BPharm, MBA, PhD

[1]School of Pharmacy and Pharmacology, University of Tasmania, Sandy Bay, Australia
[2]Tasmanian School of Medicine, University of Tasmania, Hobart, Australia

Corresponding Author:
Ivan Bindoff, BComp (Hons), PhD
School of Pharmacy and Pharmacology
University of Tasmania
1 Churchill Ave
Sandy Bay, 7005
Australia
Phone: 61 408528276
Email: Ivan.Bindoff@utas.edu.au

Abstract

Background: Numerous mobile health (mHealth) apps have been developed to support smokers attempting to quit smoking. Although these apps have been reported to be successful, only modest improvements in the quit rate have been measured. It has been proposed that efforts to improve user engagement and retention may improve the quit rate further. Owing to the high cost of smoking-related disease, it is considered worthwhile to pursue even small improvements.

Objective: The aim of this study was to test a novel smartphone app that leverages premium currency strategies developed by the mobile games industry in an attempt to improve engagement and retention with a smoking cessation intervention.

Methods: We designed and developed a smoking cessation app called "Quittr" in line with previously developed smoking cessation mHealth apps. In addition to this established framework, we added a stand-alone fully featured city-building clicker-style game called "Tappy Town," and a premium virtual currency called "QuitCoins." The user earns QuitCoins for using the app in a way that contributes positively toward their quit attempt, and they can redeem these coins in Tappy Town for bonuses. To establish whether these features improved engagement and retention, we ran a 5-month randomized controlled trial where the intervention group had the full app with the extra games features, while the control group had the standard app only. Recruitment was performed via web-based advertising. Participants (N=175) had no direct contact with the researchers or other support staff.

Results: No significant differences in terms of engagement, retention, or smoking outcomes were found between the control and intervention groups. However, survey data indicated that the majority of the participants valued Tappy Town (10/17, 59%) and the QuitCoins rewards system (13/17, 77%). Usage data also suggested that Tappy Town was widely played and was generally appealing to users (mean total time spent in app, control group: 797 seconds vs intervention group: 3502 seconds, $P<.001$). Analysis of the results suggests that users in the intervention group may have been negatively affected by the aspects of the chosen design, and some theories were explored to explain this unexpected outcome.

Conclusions: Although the novel features of the Quittr app failed to improve the key outcomes measured in this study, there were enough positive indications to warrant further exploration of the concept. Additional research will be required to identify and correct any design flaws that may have adversely affected our participants before a follow-up study can be completed.

Trial Registration: Australian and New Zealand Clinical Trials Register ACTRN12617000491369; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=372661&isReview=true

XSL•FO
RenderX

KEYWORDS

smoking; cessation; Quittr; engagement; retention; churn; cigarette; mHealth; game

Introduction

Cigarette smoking is a major preventable cause of death, disease, and financial burden. Quitting smoking is the best thing an individual can do to reduce their risk of developing smoking-related diseases [1]. Providing smoking-specific behavioral support (eg, education, advice, assistance with goal-setting) is known to improve the likelihood of succeeding in a quit attempt [2], but effectively delivering such support and keeping the individual engaged with the process for a sufficient duration is difficult.

Mobile health (mHealth) apps and web-based support tools have typically only produced smoking cessation outcomes in the order of ~9.5% success rates. Although this is a clear improvement over unassisted quit attempts and represents potentially significant savings in avoided disease burden, the outcome is still modest [3]. The difference between the known potential of this content and the actual outcomes achieved via mHealth apps may be attributed to lack of user engagement and retention [4].

This problem is similar to that faced by mobile game designers, when attempting to produce profitable mobile games that collect revenue from in-app purchases and advertising—the so called "free-to-play" games. Games are expensive to produce. However, in-game advertising typically only returns up to a few cents per ad impression. Therefore, it is critical not only to reach a very large audience but also to ensure that the audience is retained over a relatively long time span. Similarly, with in-app purchases, it is generally necessary to ensure that users are thoroughly engaged with the game experience before they are enticed with paid content that they would believe would further enhance their experience. Recognizing these similarities, we developed a smoking cessation mHealth app, which leveraged tactics from the mobile games industry to drive engagement and retention. Instead of using these tactics to drive ad impressions or in-app purchases, our objective was to use them to encourage user engagement with behavioral support content and stay on track with their quit attempt for a longer time span. We called this app Quittr. The design process and philosophy that we used has been described in our earlier paper [5].

Previous smoking cessation mHealth apps included gamification elements to encourage user engagement and provide a sense of achievement and progression, and many included mini-games to keep the smokers' hands occupied and to distract them from cravings [4,6-9]. Quittr also includes these elements as standard. However, to our knowledge, Quittr is the first such app to trial the inclusion of a fully-featured standalone game and a premium currency reward system.

For this pilot study, our aim was to determine whether these novel features added to the Quittr app could improve engagement with the smoking cessation support content and retention of the user making the quit attempt. The secondary outcome was to determine whether there was any effect on smoking cessation.

Methods

App Design

The Quittr app was broadly based upon existing smoking cessation apps (particularly SmokeFree-28 [4]) and standard behavioral content from other support tools. We included the following base features: (1) recording of daily cigarette usage, (2) a survey tool for collecting data about the user, (3) a dashboard with visualized statistics about changes to the user's health and financial position, as calculated based on user-provided data, (4) an achievement system, to reward using the app and smoke-free periods, (5) an information toolbox, with over 12,000 words of quit smoking information and advice, broken into categories/sections, and simple multiple choice quizzes at the end of each section, (6) a distraction mini-game called "Hidden Object Game," and (7) a daily notification system to remind the user to check-in and report their cigarette usage and to remind them to complete the exit survey (when appropriate).

In addition to these core features, we also added a fully featured game called "Tappy Town." This game was designed to be useful to distract the user from cigarette cravings as well as to provide a long-term incentive to engage with the broader Quittr app. We chose a simple city builder style game since it was considered both achievable within budget/time frame and to be broadly appealing to potential users across the gender and age spectrum. We also intentionally avoided any potential smoking triggers by avoiding the topic of smoking within the game.

The general design mantra throughout development was that the user could not do anything to sabotage or adversely affect their gameplay experience and could thus play without any stress. However, they could plan and optimize their play to progress faster, rewarding effort and hopefully helping to make the game more engaging.

The game was designed as a "clicker-style" game, wherein the player would gradually accrue resources and then spend these resources to build more structures, which would in turn help them to generate resources faster. As they progress, the player unlocks more expensive and better structures, and the economy generally follows an exponential growth curve. This game model has proven in recent years to be highly compelling, at least for a few days of play, as the player chases down bigger and bigger numbers and unlocks grander structures with a strong sense of progression [10].

We made some modifications from the traditional model to help the smoker to quit. As well as the city passively earning currency, the game rewarded repeat interaction from the players' hands by asking them to tap coins, which would periodically generate above their buildings to collect bonuses, as can be seen in Figure 1C.

Figure 1. Examples of various screens in the Quittr app. A. Main dashboard; B. Earning QuitCoins after playing a distraction mini-game; C: Placing a new building in the Tappy Town game.
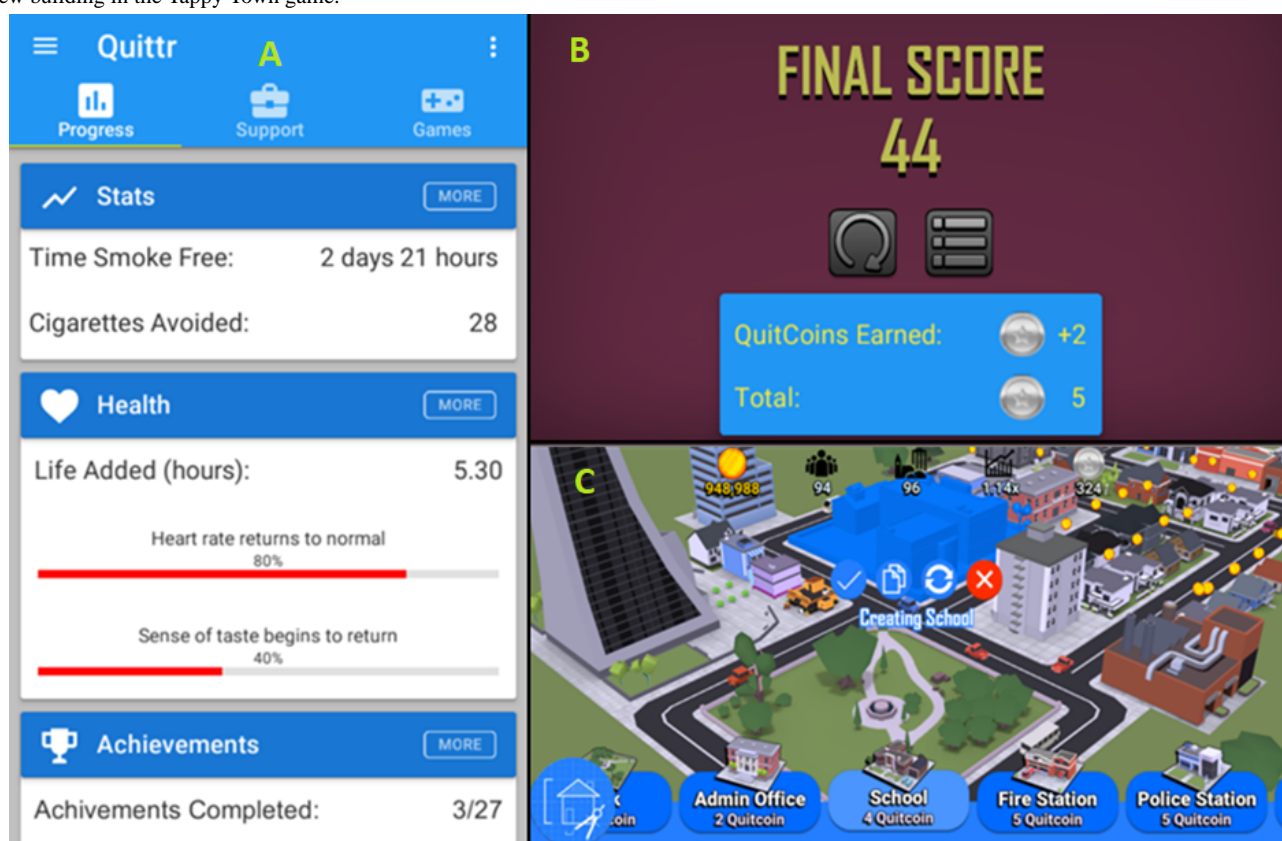
We also tuned the game to encourage playing in short sessions at regular intervals throughout the day, which was roughly aligned with the frequency with which we would expect someone who is quitting smoking to experience cravings. Our intention here was that whenever someone was experiencing withdrawals, they could log in to the game and productively play a 2-5-minute session that would help them through the worst of their symptoms, without excessive interruption of their real life.

Additionally, users could earn QuitCoins, a type of premium currency [11] by using the broader Quittr app. They could then redeem these QuitCoins in the Tappy Town game to purchase powerful bonus structures, which were otherwise not possible to be purchased. Users were rewarded with QuitCoins via the achievements system when they (1) had a breakthrough smoke-free period, (2) avoided a certain number of cigarettes, (3) checked their statistics for the first time that day, (4) engaged with the information toolbox content for set periods of time, (5) successfully completed an information toolbox quiz, (6) played Tappy Town for the first time in a day, (7) played the Hidden Object Game for the first time in a day, or (8) set savings goals.

This mechanism was intended to facilitate the transfer of the player from the game back to the helpful quit smoking content and vice versa, and it was intended that the player would be motivated to engage with the quit-smoking content more closely than they might have otherwise, in order to advance in the Tappy Town game. However, little to no explicit messaging about these features was provided to the user. They were instead expected to experience this effect organically without any of

the "nag screens" that would typically accompany this type of design in free-to-play mobile games. This was done to avoid adding potential stressors.

Experimental Design

To test our innovation, we designed a randomized controlled trial with 2 groups: control and intervention (Multimedia Appendix 1). The intervention group had access to the full Quittr experience, including our games-based innovations and the QuitCoins premium currency incentivization system. The control group also had access to Quittr, but they did not have access to the Tappy Town game and could not earn QuitCoins. The protocol was registered with the Australian and New Zealand Clinical Trials Register [ACTRN12617000491369] and ethics approval was granted by the Tasmanian Health and Medical Human Research Ethics Committee [H0016506]. Participants were eligible if they were current smokers interested in starting a quit attempt, with a suitable smartphone device, aged 15 years and over, and spoke English. Allocation was done in a pseudo-random manner that was not known to participants, to ensure roughly equal group sizes. At semifrequent periods throughout the day, an algorithm was run to determine which group currently had more members and at which point the group with the fewest members would become favored for new registrations.

Recruitment

Recruitment was open from December 12, 2017 through March 9, 2018 and was done primarily through posts on web-based forums and targeted advertising through Facebook (Multimedia Appendix 2). Participants were able to trial the app in testing

mode and start a quit attempt at any time. Once they started a quit attempt, we tracked their data over a 28-day period. Although a few users elected to abandon 1 quit attempt and start a new quit attempt during the trial period, each quit attempt was evaluated individually, with the participant's longest attempt being used for the final analysis.

Data Collection

We logged all participant activity within the app, which was automatically uploaded to our server when internet access was detected via Wi-Fi. To answer the research questions on engagement and retention, we analyzed how frequently the participants were using the app, time spent using the various features of the app, and the time when usage occurred. We also administered basic surveys from within the app: an entry survey at the commencement of a quit attempt and an exit survey when the quit attempt was finished (28-day window completed) or abandoned (user-indicated) (Multimedia Appendix 3).

Analysis

To determine whether the groups were equivalent in terms of participant characteristics at baseline, we used chi-square tests to compare categorical variables and two-sided independent samples t tests to compare means. When comparing the groups and outcomes at the end of the trial, we used means and two-sided independent samples t tests wherein data were

parametric and medians and two-sided Wilcoxon rank sum tests wherein data were nonparametric. To explore engagement, we compared the time spent using the various features of the app between groups and analyzed retention by comparing the last day of usage between the groups. An equivalent test was also run to determine whether 1 group had a greater success at quitting smoking by comparing the longest smoke-free period of the members of each group. We used a chi-square test to determine differences in user perceptions indicated by the exit survey results and descriptive statistics to explore questions that were unique to 1 group only.

Results

Participant Characteristics

During the 4-month recruitment window, we recruited 182 participants. Unfortunately, an issue was found in the first week, which caused 6 participants to have their group incorrectly coded; thus, their results were discarded, leaving us with 176 participants and 209 quit attempts. An additional user was excluded for submitting unrealistic results, leaving us with 87 participants in the control group and 88 in the intervention group (N=175), although 23 did not submit survey responses beyond age and gender (12 in control and 11 in intervention). The 2 groups were well matched, as shown in Table 1.

Table 1. Comparisons of the demographic and smoking characteristics, based on entry survey data, between the control and intervention groups (N=175).

Demographic/smoking characteristics	Control group (n=87)	Intervention group (n=88)	χ^2 (df)	P value
Age (years), mean (SD)	39.5 (12.75)	40.5 (15.48)	N/A[a]	.36
Gender, n (%)			0.73	.24
Female	44 (51)	44 (50)		
Male	43 (49)	44 (50)		
How soon after waking do you smoke your first cigarette? (minutes), n (%)[b]			3.62 (4)	.46
<5	24 (32)	24 (31)		
6-15	18 (24)	20 (26)		
16-30	17 (23)	11 (14)		
31-60	11 (15)	11 (14)		
>60	5 (7)	11 (14)		
How many cigarettes do you smoke per day on average? mean (SD)	20.27 (20.36)	19.53 (22.48)	N/A	.83
Are you intending to use nicotine replacement therapy? n (%)[b]			0.21 (1)	.65
No	46 (61)	50 (65)		
Yes	29 (39)	27 (35)		
Are you intending to use medication to manage cravings/withdrawals? n (%)[b]			0.20 (1)	.66
No	67 (89)	67 (87)		
Yes	8 (11)	10 (13)		

[a]N/A: not applicable.

[b]For these categories, only 75 participants in the control group and 77 participants in the intervention group provided responses.

Participant Engagement With Quittr

Participants in the intervention group did post higher overall app usage statistics (mean time, intervention group: 3502 seconds vs control group: 797 seconds, $P<.001$) (Table 2). However, it appears that this extra time was almost entirely spent using the Tappy Town game. There were no other significant differences in the amount of time spent between the control and intervention groups, although the intervention group did trend toward spending additional time using the information toolbox (mean 189 seconds [intervention] versus 154 seconds [control]). The finance stats were close to statistical significance, but with only a little over 3 seconds separating them was not considered meaningful.

Table 2. Time spent by both the groups using the various features of Quittr.

Features of Quittr	Control group, average time, (seconds)	Intervention group, average time, (seconds)	Mean difference (seconds)	Standard error difference (seconds)	P value
Main dashboard	312.64	364.42	–51.78	74.50	.22
Information toolbox	154.39	188.79	–34.40	52.87	.11
Health stats	33.54	32.69	0.85	12.82	.84
Smoking stats	27.91	17.65	10.26	8.89	.19
Finance stats	6.96	10.65	–3.69	3.70	.06
Achievements	20.70	25.95	–5.25	9.79	.22
Tappy Town game	0.00	2651.78	–2651.78	709.97	N/A[a]
Hidden object game	132.81	99.89	32.92	32.79	.65
Total time in app	796.65	3501.86	–2705.21	816.73	*<.001* [b]
Total time (excluding Tappy Town)	796.65	850.08	–53.43	167.43	.42

[a]N/A: not applicable.

[b]This result is significant at $P<.05$.

It was anticipated that the premium currency approach might be compelling only for a minority of users; therefore, the effect may not be demonstrable through comparison of the means. It is possible to visually inspect the outliers by plotting the time spent engaging with Tappy Town against the time spent engaging with the behavioral support content in the app (Figure 2).

Figure 2. Time spent engaging with behavioral support content versus time spent in Tappy Town.
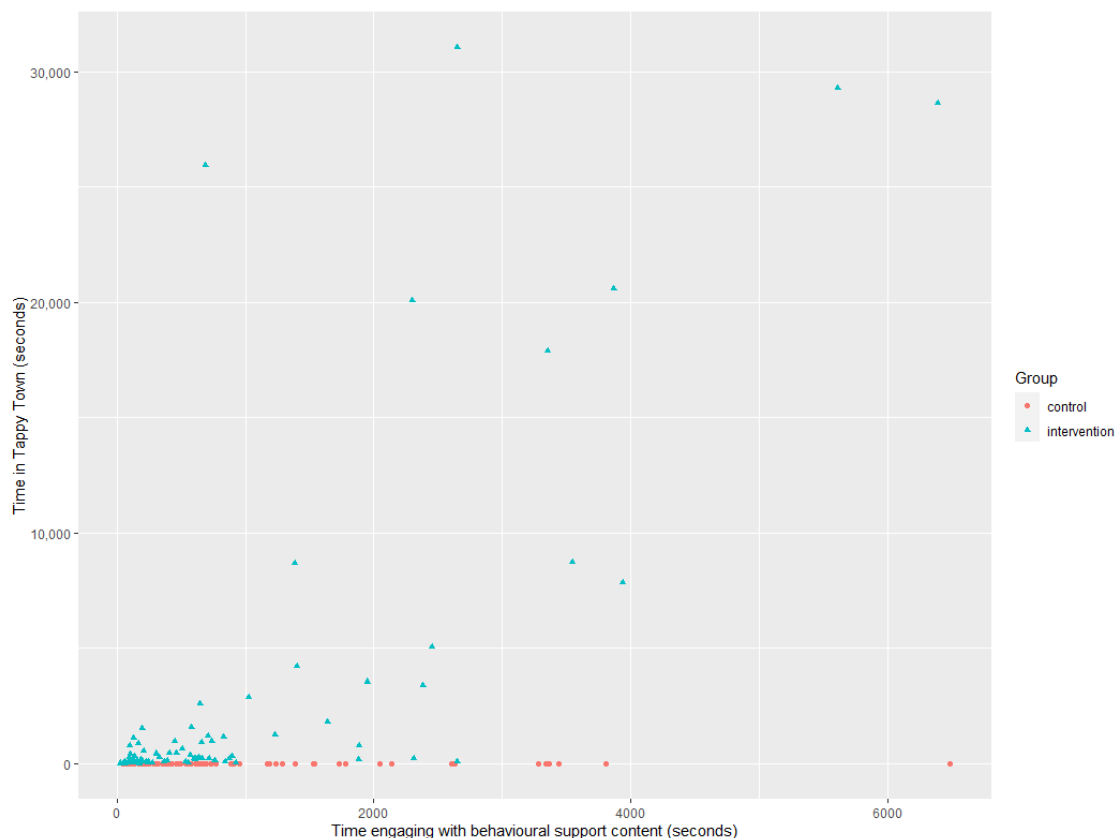
Participant Retention

The mean (SD) retention time of our participants was 6.67 (8.88) days. There was no statistically significant difference in the median number of days of app use between the control and intervention groups (median 2, IQR=0-12 vs median 1.5, IQR=0-9 days, respectively; P=.17).

Smoking Status of the Participants

Only 7 of our 175 participants (4.0%) recorded 28 days smoke-free, with a mean (SD) smoke-free period of 4.80 (7.40) days. This result contrasts starkly with the results of the SmokeFree-28 app upon which Quittr was based, which saw 18.9% of their participants successfully achieve 28 days smoke-free. There was no statistically significant difference in the maximum achieved smoke-free period between the control and intervention groups (median 1, IQR=1-3.50 vs median 1, IQR=1-5 days, respectively; P=.59).

Survey Responses

A total of 27 (31%) participants in the control group (n=87) submitted exit survey responses compared to only 17 (19%) for the intervention group (n=88) (P=.08). Although not reaching significance, this apparent disparity in the submissions was an unexpected outcome. An additional two-sided t test was performed to determine if the baseline smoking profile of the participants who submitted exit survey results was different from those who did not. Although the mean cigarettes per day statistic did appear slightly lower, this did not reach statistical significance (20.9 cigarettes per day vs 16.6 cigarettes per day, respectively; P=.10).

General Responses

Of the exit survey questions that were asked to both groups and where both surveys were completed by the user, there were no significant differences between the groups (Table 3).

Table 3. Comparison of the responses to the exit survey questions by both groups.

Response	Control group (n=27), n (%)	Intervention group (n=17), n (%)	χ^2 *(df)*	*P* value
The statistics on the Dashboard helped me in my quit attempt			0.99 (4)	.91
Strongly disagree	1 (4)	1 (6)		
Disagree	1 (4)	0 (0)		
Neutral	6 (22)	4 (24)		
Agree	11 (41)	8 (47)		
Strongly agree	8 (30)	4 (24)		
The educational content in the Information Toolbox helped me in my quit attempt			1.76 (4)	.78
Strongly disagree	1 (4)	0 (0)		
Disagree	1 (4)	0 (0)		
Neutral	8 (30)	7 (41)		
Agree	11 (41)	6 (35)		
Strongly agree	6 (22)	4 (24)		
The games helped me in my quit attempt			3.77 (4)	.44
Strongly disagree	2 (7)	1 (6)		
Disagree	2 (7)	1 (6)		
Neutral	14 (52)	9 (53)		
Agree	5 (19)	6 (35)		
Strongly agree	4 (15)	0 (0)		
The games helped me avoid cravings			4.36 (4)	.36
Strongly disagree	3 (11)	1 (6)		
Disagree	1 (4)	2 (12)		
Neutral	16 (59)	9 (53)		
Agree	4 (15)	5 (29)		
Strongly agree	3 (11)	0 (0)		
How often were you truthful when reporting your daily cigarette intake?			3.6 (3)	.31
Never	0 (0)	0 (0)		
Rarely	1 (4)	1 (6)		
Sometimes	4 (15)	0 (0)		
Very often	2 (7)	3 (18)		
Always	20 (74)	13 (77)		

Participant Perceptions of the Games-Based Features of Quittr

The intervention group was asked several questions to ascertain their perceived responses to the games-based features, as shown in Table 4. This group generally showed enthusiasm toward the games features, with the majority of users agreeing that the Tappy Town game was engaging, QuitCoins rewards were enjoyable, and they wanted to earn QuitCoins and spend them in Tappy Town.

Table 4. Participants' perceptions of the games-based features of Quittr (n=17).

Perceptions, responses	n (%), Value
I found the Tappy Town game engaging	
Strongly disagree	1 (6)
Disagree	0 (0)
Neutral	6 (35)
Agree	9 (53)
Strongly agree	1 (6)
I enjoyed earning QuitCoins rewards	
Strongly disagree	1 (6)
Disagree	1 (6)
Neutral	2 (12)
Agree	11 (65)
Strongly agree	2 (12)
I wanted to earn QuitCoins rewards so I could use them in Tappy Town	
Strongly disagree	1 (6)
Disagree	2 (12)
Neutral	5 (29)
Agree	8 (47)
Strongly agree	1 (6)

Smoking Statistics

Amongst those participants who completed the exit survey, users reported a median daily cigarette reduction of 7 (IQR 0-15) and a weekly median cigarette cost (reported in participant's own local currency, USD or AUD) reduction of $0 (IQR $0-$19.5). When we compared the groups, these statistics *favored* the control group. The median reduction in daily cigarettes was 12 (IQR 8.5-20) for the control group and 0 (IQR 0-4.5) for the intervention group ($P<.001$), while the median weekly cigarette cost reduction was $8.50 (IQR $0-$46.25) for the control group and $0 (IQR $13-$15) for the intervention group ($P=.069$).

Discussion

Principal Findings

Our games-based innovations did successfully increase participant engagement. However, at least with our implementation, their additional engagement was limited to the game itself with little of the desired flow-on effect into other aspects of the app. Nevertheless, it is somewhat encouraging that significant play time was invested into the Tappy Town game, since one of the key concerns with this study was that we would not be able to create a compelling game experience with the time and budget we had available. Unfortunately, the results show that our efforts to incentivize the participants to engage with the broader app were not effective. Similarly, our strategies were not effective in improving the retention of users. However, it is not yet clear whether these failings were due to fundamental flaws with the approach or flaws with our chosen design and implementation.

These relatively poor outcomes become perplexing when we examine the user perceptions from the exit survey, which showed that the majority of users in the game group valued the games features we added and felt that they were incentivized by them. Given this unexpected outcome and the benefit of hindsight, it would have been useful to include survey questions to explore user-perceived engagement in the broader Quittr app, so as to provide insight into whether the user *feels* (dis)engaged by a given aspect of the app and why and to allow more detailed comparisons of engagement between the groups. Nevertheless, the results of this survey combined with the significant number of play-minutes recorded in Tappy Town suggest that our hypothesis may have been undermined by other factors; therefore, significant effort was expended in an attempt to understand what these may be.

Learnings

Perhaps the most surprising result of the exit survey was that 27 participants in the control group submitted results compared to only 17 in the intervention group and that among these users, the control group achieved significantly better daily cigarette reductions. In effect, what we observed here was that the intervention group had noticeably *worse* retention—the users were less likely to receive or act upon the notification to complete the 28-day exit survey. This prompted significant exploration and discussion among the project team, especially when considered alongside the smoking statistics, which unexpectedly favored the control group.

XSL•FO
RenderX

Considering that the intervention group had access to all the same features and support as the control group, we must consider how this might have occurred. One possible explanation is that the Tappy Town game actively stole users' attention away from the smoking cessation features of the Quittr app and interfered with the success of their quit attempt. However, users in the intervention group spent just as much time engaging with the smoking cessation content of the app as the control group did; so, this explanation seems unlikely. Another possibility is that the poorer outcome was simply due to chance—this is quite possible, since our sample size was relatively modest and the only statistically significant difference was in the very specific daily cigarette reduction outcome—this could simply be a type I statistical error.

Yet another possibility is that when users grew tired of either the Tappy Town game or the smoking cessation content, they may have uninstalled the entire Quittr app. In contrast, participants in the control group could not become tired of the Tappy Town game; thus, they were less likely to uninstall—this explanation aligns to some degree with activity engagement theory, where experiments have demonstrated that students being given 2 unrelated tasks have their subsequent engagement diminished for both tasks [12].

If we explore theories regarding user retention and engagement through a technology acceptance lens, some other possible explanations arise. It is possible that the addition of the Tappy Town game made the Quittr app feel *less easy to use,* or in other words, required *higher effort* [13,14], or perhaps even made the educational content in the app seem *less credible* [14-16] via its association with a relatively nonserious game experience. Additional research would be required to determine from participants which, if any, of these factors may have come into play and potentially explore strategies to ameliorate the effect. Another likely possibility is that we failed to adequately capture one of the key motivational messages of the SmokeFree-28 app, which quite intensively conveyed the goal of going 28 days smoke-free as part of its PRIME motivational theory-based design. As the authors stated, "the core of SmokeFree-28 involves setting a highly salient target of becoming 28 days smoke-free and monitoring progress towards that target using the app" [4]. Although Quittr did set an equivalent target, it was not as extensively highlighted and focused as in SmokeFree-28; rather, it was just one of many metrics reported on the dashboard—one of the many achievements that could be attained. This may at least partially explain why neither group in our trial managed to achieve equivalent outcomes to the SmokeFree-28 trial (18.9% of the SmokeFree-28 participants successfully reported 28 days of smoke-free status compared to Quittr's control group at 6%).

Analysis of the SmokeFree-28 trial compared quit success against the Smoking Toolkit Study as a baseline, where 15% of the smokers without any app support were able to successfully quit for 28 days. This baseline still contrasts starkly with our figures, which suggests there may also be fundamental differences in the participants that our respective trials were able to recruit. Perhaps the most logical explanation for this is that the Quittr study was advertised to smokers via social media and targeted web-based advertising, whereas the SmokeFree-28

study appears to have achieved organic participation through user-initiated search and download—this may have created selection bias, with their participants being more self-motivated to quit, whereas our participants were perhaps biased toward strongly identifying as a smoker in their social media presence, and thus becoming the recipient of targeted advertising. This theory is supported when examining our participant profile and recognizing that our participants smoked an average of 20 cigarettes per day and overwhelmingly smoked their first daily cigarette within 30 minutes of waking—this contrasts with the SmokeFree-28 data where ~65% of the participants smoked fewer than 19 cigarettes per day. It appears that our sample was biased toward heavy smokers, which may have made it harder to achieve the 28-day cessation outcome.

There is a final possibility to consider. The exit survey was prompted by a phone notification that occurred after a period of participant inactivity or after the 28-day study period was complete. This feature worked as intended; however, it could be thwarted by users if they explicitly turned off notifications for the Quittr app or uninstalled the Quittr app. It is possible that participants in the intervention group may have been more likely to turn off notifications or uninstall the app. The app would send notifications once daily at 8 PM. There were 3 reminders, which would fire if appropriate: one to remind the user to log their cigarette usage, which would always fire, one to prompt them to complete the exit survey if the user had been inactive for an extended period or had reached 28 days of participation, and a third that only applied to intervention group users, which would prompt the user to spend excess resources in Tappy Town if their accrued resources were over a given threshold. Consequently, most participants in the intervention group would receive 2 simultaneous notifications each night, whereas most control participants would receive only 1. We suspect this "spammy" activity may have prompted some of them to turn off the notifications or uninstall the app.

Unfortunately, we have no way of knowing whether they did turn off notifications or uninstall the app, or when. However, it may also explain why we saw no positive intervention effect on retention in general—if participants in the intervention group were more likely to turn off notifications, then they were missing out on a critical aspect of our user retention strategy: daily reminders to log in, log cigarette usage, and generally engage with the content of the app. In light of this, it is appropriate that future studies should examine their data collection strategies with care and ideally include features to log whether notifications for the app were disabled and attempt to track when the app is uninstalled.

There is another result that suggests there may be potential for design improvement. Only 9 of the 17 respondents (53%) reported wanting to earn QuitCoins to spend them in Tappy Town compared to 13 (77%) who said they enjoyed earning QuitCoins. This difference may simply be that some users were not motivated by the Tappy Town game but did like the feeling of achievement that the QuitCoins rewards gave them. However, on reflection, we suspect that more should have been done to make the link between earning QuitCoins and spending them in Tappy Town. There were no explicit messages or reminders for the user to earn and spend QuitCoins, and at very few points,

it was explained to the user that QuitCoins could be spent in Tappy Town to build powerful structures. This is unlike most mobile games that use a similar premium currency model.

Conclusions and Further Work

Although our pilot study achieved poor results in terms of both the primary and secondary outcomes, it has nevertheless provided some useful insights and some potentially encouraging preliminary data. Survey responses suggest that the hypothesis "games features can be used to encourage engagement and retention in a quit smoking app" has some credibility on face value at least, and the significant play-time invested by participants into the Tappy Town game suggests that the game itself was acceptable. It appears as if there may be other factors at play that adversely affected the outcomes of the intervention group, although perhaps the game itself negatively affected retention through one of the discussed theoretical mechanisms or a mechanism still unknown.

Without these adverse factors in play, would a positive outcome have been achieved? To answer that question, it would be necessary to categorically identify and fix the issues before rerunning the experiment. Although we have identified some potential issues, this study was not designed for this purpose; therefore, there may be other issues yet unidentified. We would recommend commissioning first a smaller scale usability analysis with observed and qualitatively studied users. This study should identify potential pain points and points of conflict that could negatively impact the outcomes of the intervention, which can then be remedied before undertaking any further quantitative study.

Conflicts of Interest

SGF provides consulting services to numerous companies on matters relating to smoking cessation and harm reduction. The other authors report no conflicts of interest.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1153 KB - games_v8i4e23734_app1.pdf]

Multimedia Appendix 2
Media release and Twitter advertisements.
[DOCX File , 410 KB - games_v8i4e23734_app2.docx]

Multimedia Appendix 3
Entry and exit surveys.
[DOCX File , 14 KB - games_v8i4e23734_app3.docx]

References

1. Critchley JA, Capewell S. Mortality risk reduction associated with smoking cessation in patients with coronary heart disease: a systematic review. JAMA 2003 Jul 02;290(1):86-97. [doi: 10.1001/jama.290.1.86] [Medline: 12837716]
2. Michie S, Churchill S, West R. Identifying evidence-based competences required to deliver behavioural support for smoking cessation. Ann Behav Med 2011 Feb;41(1):59-70. [doi: 10.1007/s12160-010-9235-z] [Medline: 20936389]
3. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2012 Nov 14;11:CD006611. [doi: 10.1002/14651858.CD006611.pub3] [Medline: 23152238]
4. Ubhi HK, Michie S, Kotz D, Wong WC, West R. A mobile app to aid smoking cessation: preliminary evaluation of SmokeFree28. J Med Internet Res 2015 Jan 16;17(1):e17 [FREE Full text] [doi: 10.2196/jmir.3479] [Medline: 25596170]
5. Bindoff I, de Salas K, Peterson G, Ling T, Lewis I, Wells L, et al. Quittr: The Design of a Video Game to Support Smoking Cessation. JMIR Serious Games 2016 Dec 01;4(2):e19 [FREE Full text] [doi: 10.2196/games.6258] [Medline: 27908844]
6. Krebs P, Burkhalter JE, Snow B, Fiske J, Ostroff JS. Development and Alpha Testing of QuitIT: An Interactive Video Game to Enhance Skills for Coping With Smoking Urges. JMIR Res Protoc 2013 Sep 11;2(2):e35 [FREE Full text] [doi: 10.2196/resprot.2416] [Medline: 24025236]

7.   El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game On? Smoking Cessation Through the Gamification of mHealth: A Longitudinal Qualitative Study. JMIR Serious Games 2016 Oct 24;4(2):e18 [FREE Full text] [doi: 10.2196/games.5678] [Medline: 27777216]

8.   Marin-Gomez FX, Garcia-Moreno Marchán R, Mayos-Fernandez A, Flores-Mateo G, Granado-Font E, Barrera Uriarte ML, et al. Exploring Efficacy of a Serious Game (Tobbstop) for Smoking Cessation During Pregnancy: Randomized Controlled Trial. JMIR Serious Games 2019 Mar 27;7(1):e12835 [FREE Full text] [doi: 10.2196/12835] [Medline: 30916655]

9.   Edwards EA, Caton H, Lumsden J, Rivas C, Steed L, Pirunsarn Y, et al. Creating a Theoretically Grounded, Gamified Health App: Lessons From Developing the Cigbreak Smoking Cessation Mobile Phone Game. JMIR Serious Games 2018 Nov 29;6(4):e10252 [FREE Full text] [doi: 10.2196/10252] [Medline: 30497994]

10.   Alharthi S, Alsaedi O, Toups Z, Tanenbaum J, Hammer J. Playing to wait: A taxonomy of idle games. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI Conference on Human Factors in Computing Systems; 2018; Montreal p. 1-15. [doi: 10.1145/3173574.3174195]

11.   Evans E. The economics of free: Freemium games, branding and the impatience economy. Convergence 2016 Jul 08;22(6):563-580. [doi: 10.1177/1354856514567052]

12.   Higgins ET, Lee J, Kwon J, Trope Y. When combining intrinsic motivations undermines interest: a test of activity engagement theory. J Pers Soc Psychol 1995 May;68(5):749-767. [doi: 10.1037//0022-3514.68.5.749] [Medline: 7776180]

13.   Venkatesh, Morris, Davis, Davis. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]

14.   Yoganathan D, Kajanan S. What drives fitness apps usage? An empirical evaluation. In: International Working Conference on Transfer and Diffusion of IT.: Springer; 2014 Presented at: International Working Conference on Transfer and Diffusion of IT; 2014; Aalborg p. 179-196. [doi: 10.1007/978-3-662-43459-8_12]

15.   Pavlou PA. Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model. International Journal of Electronic Commerce 2014 Dec 23;7(3):101-134. [doi: 10.1080/10864415.2003.11044275]

16.   Lanseng EJ, Andreassen TW. Electronic healthcare: a study of people's readiness and attitude toward performing self-diagnosis. Int J of Service Industry Mgmt 2007 Aug 14;18(4):394-417. [doi: 10.1108/09564230710778155]

Abbreviations

mHealth: mobile health

Original Paper

Use of Virtual Reality to Assess Dynamic Posturography and Sensory Organization: Instrument Validation Study

Matthew William Wittstein[1*], BSc, MSc, PhD; Anthony Crider[2*], PhD; Samantha Mastrocola[1*], BSc; Mariana Guerena Gonzalez[1*], BSc

[1]Department of Exercise Science, Elon University, Elon, NC, United States

[2]Department of Physics, Elon University, Elon, NC, United States

[*]all authors contributed equally

Corresponding Author:

Matthew William Wittstein, BSc, MSc, PhD
Department of Exercise Science
Elon University
2525 Campus Box
Elon, NC, 27244
United States
Phone: 1 336 278 6693
Email: mwittstein@elon.edu

Abstract

Background: The Equitest system (Neurocom) is a computerized dynamic posturography device used by health care providers and clinical researchers to safely test an individual's postural control. While the Equitest system has evaluative and rehabilitative value, it may be limited owing to its cost, lack of portability, and reliance on only sagittal plane movements. Virtual reality (VR) provides an opportunity to reduce these limitations by providing more mobile and cost-effective tools while also observing a wider array of postural characteristics.

Objective: This study aimed to test the plausibility of using VR as a feasible alternative to the Equitest system for conducting a sensory organization test.

Methods: A convenience sample of 20 college-aged healthy individuals participated in the study. Participants completed the sensory organization test using the Equitest system as well as using a VR environment while standing atop a force plate (Bertec Inc). The Equitest system measures the equilibrium index. During VR trials, the estimated equilibrium index, 95% ellipse area, path length, and anterior-posterior detrended fluctuation analysis scaling exponent alpha were calculated from center of pressure data. Pearson correlation coefficients were used to assess the relationship between the equilibrium index and center of pressure–derived balance measures. Intraclass correlations for absolute agreement and consistency were calculated to compare the equilibrium index and estimated equilibrium index.

Results: Intraclass correlations demonstrated moderate consistency and absolute agreement ($0.5 <$ intraclass correlation coefficient < 0.75) between the equilibrium index and estimated equilibrium index from the Equitest and VR sensory organization test (SOT), respectively, in four of six tested conditions. Additionally, weak to moderate correlations between force plate measurements and the equilibrium index were noted in several of the conditions.

Conclusions: This research demonstrated the plausibility of using VR as an alternative method to conduct the SOT. Ongoing development and testing of virtual environments are necessary before employing the technology as a replacement to current clinical tests.

Introduction

The Equitest system (Neurocom) is a computerized dynamic posturography device used by health care providers and clinical researchers to safely test an individual's postural control. Implementing the sensory organization test (SOT) using the Equitest system requires individuals to process and integrate cues from the visual, vestibular, and proprioceptive systems. This test provides clinicians and researchers with an equilibrium score for each tested condition, a sensory analysis score, a strategy analysis, and a center of gravity (COG) alignment. While the Equitest system has evaluative and rehabilitative value, it may be limited owing to its cost and lack of portability. Moreover, the performance variables provided by the Equitest system are limited, representing gross outcome measures derived only from sagittal plane movement dynamics [1]. Recent advances in technology provide opportunities to reduce these limitations by providing more mobile and cost-effective tools while also observing a wider array of postural characteristics. The purpose of this research was to evaluate the validity of using virtual reality (VR) and a force plate as an alternative to the Equitest system.

The SOT has been the dominant clinical test to assess sensory integration in the context of postural control for neurologic disorders and deficits. With the wide use of clinical dynamic posturography over the last 30 years, the Equitest system has become widely accepted as the gold standard to assess postural stability and balance in several populations (eg, children, aging adults, and military personnel) and clinical groups (eg, those with concussion, vertigo, Parkinson disease, and Alzheimer disease). By systematically disrupting the visual and somatosensory information available to an individual, it is possible to distinguish someone's reliance on the following three major sensory systems during balance tasks: the visual, somatosensory, and vestibular systems. Conveniently, the Equitest system provides an equilibrium score (indicating how little participants swayed) during each test, as well as a sensory analysis score (indicating how much they relied on each system) and strategy analysis (indicating the hip versus ankle strategy) for the battery of conditions.

While the Equitest system provides a quick evaluative tool for clinicians and researchers, it is not without limitations. First, these outcome measures are derived solely from sagittal plane movements and may not reflect a complete assessment of an individual's postural control. Second, the costs associated with the Equitest system may limit its availability in underserved communities or during times immediately following an injury (such as a sports concussion). As an alternative to the Equitest system, it may be possible to combine more recent technologies, that is, portable force plates and VR, to ameliorate these

drawbacks. When these technologies are combined, they greatly reduce the cost for a clinician to own testing equipment, as well as offer the opportunity to have a portable solution that could be taken into the field. Moreover, portable force plates present the possibility to record and assess a wider range of data, such as medial-lateral dynamics, and customize the outcomes to specific clinical goals. Likewise, VR headsets have continued to improve in quality and decrease in cost, and continued developments may lead to the ability to accurately track movements in VR without additional hardware components such as force plates.

In keeping up with technological advancements, it is important to determine how new technologies can measure up to the "gold standards" they will eventually replace. Currently, VR is approaching this standard and is consistently shown to be a valuable tool to conduct postural and motor control research. Previous research has found no difference between static balance in a physical environment versus a virtual environment [2]. Additionally, several scholars have supported the efficacy of VR for use in balance assessments in a range of clinical populations, such as those with concussion, stroke, Parkinson disease, and high age [3-7]. Continuing in this trend, a large body of research has shown positive results in using VR to enhance training and rehabilitation for balance-related dysfunction [8-11]. Overall, VR has been demonstrated to accurately assess balance in addition to providing a customizable means to enhance clinical outcomes.

The purpose of this research was to compare the Equitest system to a VR balance assessment designed to mimic the SOT in a young healthy population. It was hypothesized that the equilibrium score would demonstrate high limits of agreement between the two testing conditions, supporting VR as a viable option to decrease cost and increase the accessibility of postural assessment techniques. By illustrating the viability of VR to emulate current clinical practices, future progress can focus on improving and optimizing the implementation of VR in clinical standards of care and applications to more populations of interest.

Methods

Participants

A convenience sample of 20 college-aged individuals (Table 1) was recruited to participate in this study. All participants were healthy individuals with no prior history of neurological or physical injury or dysfunction. Upon arrival, participants provided informed consent. All procedures were approved by the institutional review board, and no adverse events were encountered.

Table 1. Participant demographics.

Characteristic	Male (n=7), mean (SE)	Female (n=13), mean (SE)
Age (years)	20.8 (0.4)	20.9 (0.37)
Height (m)	1.79 (0.03)	1.66 (0.02)
Weight (kg)	77.4 (5.58)	62.8 (3.33)

XSL•FO

RenderX

Experimental Design

After providing informed consent, participants completed a SOT in two blocks, using the Equitest system and using VR. Blocks of tests were counterbalanced, and conditions within blocks were randomized.

During the Equitest SOT, participants wore a harness that supported their weight in case they lost balance. Researchers helped participants into the harness so it fit comfortably and safely. The conditions during the clinical test included (1) eyes open on a stable surface, (2) eyes closed on a stable surface, (3) eyes open with a sway-referenced surround, (4) eyes open on a sway-referenced surface, (5) eyes closed on a sway-referenced surface, and (6) eyes open with both a sway-referenced surround and surface.

In the VR SOT, participants removed any glasses and wore a head-mounted display (HTC Vive, HTC). Participants adjusted the headset to ensure clarity in the virtual environment using a black screen with a textbox. To compare our VR SOT to existing SOT research performed with real machines, we created a virtual scale model of the patterns used inside of the Equitest balance system. We placed this model in the center of a white virtual testing room (10 m × 9 m in size). These models and the testing software were created using Unity 3D (v. 2018.2.10f1; Unity Technologies). Our software allowed us to test users with the following three different types of VR tracking: no tracking, head rotation tracking only, and six degrees of freedom (6DoF)

tracking (Figure 1). The "no tracking" option creates an experience where the objects viewed move with the user's head as if they are attached. The second option, which is common in first-generation VR headsets, such as the Oculus DK1 and Google Daydream, is somewhat natural until users lean in a direction that moves their torso. The last of these most closely mimics reality.

Balance was tested in the following conditions in the VR environment: in a completely dark environment, eyes open in an environment that mimics the clinical test (6DoF tracking), eyes open in an environment that mimics the surround of the clinical test and moves and rotates with the participant's head (no tracking), and eyes open in an environment that mimics the surround of the clinical test and moves forward and backward with the participant's head but does not react to head rotation (head tracking only). Each condition was completed on a stable surface and on a foam surface.

For each balance condition, in both the clinical test and the VR test, participants completed two trials of 20 seconds. The order of the trials was counterbalanced between the clinical test and VR blocks, and the order of the conditions was randomized within the clinical test and VR blocks. In total, participants completed 28 trials (six clinical testing conditions × two trials each, four VR testing conditions × two surface conditions × two trials each) of 20 seconds of stationary balance. All participants provided written consent prior to beginning the experimental protocol.

Figure 1. Effect of the head tracking condition in virtual reality on a user's view with translation or rotation of the head. 6DoF: six degrees of freedom.
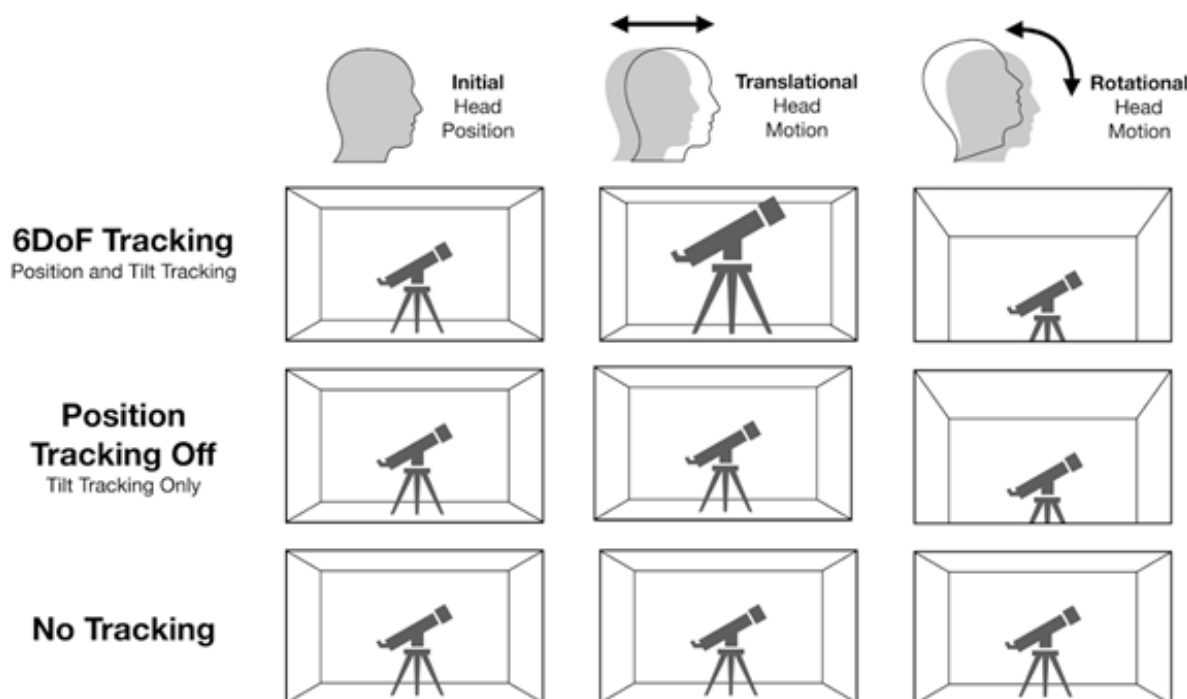
Data Reduction

The Equitest system calculated the equilibrium index (EI) during each SOT condition [12], and it represents the extent to which a participant sways forward or backward within a theoretical limit of 12.5° of displacement. If the participant has no sway,

a score of 100 would be received, and if the participant exhibits 12.5° or greater sway (combined forward and backward), a score of 0 would be received. During the VR conditions, participants completed the test on top of a portable force plate (Bertec Inc) that collected center of pressure (COP) data at 50 Hz. Custom MATLAB (Mathworks Inc) scripts were used to

detrend and filter the data (20-ms moving average filter) and subsequently calculate the estimated equilibrium index (eEI), 95% ellipse area, path length, and anterior-posterior (AP) detrended fluctuation analysis scaling exponent alpha (DFA α) from the COP data. The 95% ellipse area, path length, and AP DFA α calculations are described elsewhere and represent typical spatiotemporal characteristics of balance [13,14]. The eEI metric was derived based on the EI used by the Equitest system. To simplify this process, the forward and backward sway angles were calculated as the inverse sine function of the anterior and posterior COP displacement, respectively, divided by an estimated COG height (56% of the participant height). The first trial of each condition served as a familiarization period, and only the final trial of each condition was used for analysis. Data that were outside of three times the SD from the mean of its experimental condition were removed from the analysis. In this manner, one trial each from SOT 3 and SOT 4 was removed, along with their VR condition pair.

Statistical Analysis

To assess the relationship between EI and eEI, intraclass correlations of consistency and absolute agreement were calculated for similar conditions (Table 2). Intraclass correlation coefficient (ICC) values were interpreted as poor (<0.5), moderate (0.5-0.75), good (0.5-0.9), and excellent (>0.9) reliability [15]. Additionally, Pearson correlation coefficients were calculated to quantify the extent to which force plate measurements were associated with the EI calculated by the Equitest system within similar conditions. Correlation coefficients were interpreted as negligible (<0.3), weak (0.3-0.5), moderate (0.5-0.7), strong (0.7-0.9), or very strong (>0.9) relationships between pairs of variables [16].

Table 2. Summary of all testing conditions, their abbreviations, and the quality of visual, somatosensory, and vestibular information available in the condition.

Condition abbreviation	Equitest	Virtual reality	Information quality[a]
SOT[b] 1	Eyes opened on a stable surface	Stable virtual surround on a stable surface	Vis[c]–Som[d]–Ves[e]
SOT 2	Eyes closed on a stable surface	Blacked out environment on a stable surface	Som–Ves
SOT 3	Eyes opened with a sway-referenced surround	Head-referenced virtual surround on a stable surface	*Vis*–Som–Ves
SOT 4	Eyes opened on a sway-referenced surface	Stable virtual surround on a foam surface	Vis–*Som*–Ves
SOT 5	Eyes closed on a sway-referenced surface	Blacked out environment on a foam surface	*Som*–Ves
SOT 6	Eyes opened with a sway-referenced surround and on a sway-referenced surface	Head-referenced virtual surround on a foam surface	*Vis*–*Som*–Ves

[a]In the column, normal text indicates accurate and italic text indicates inaccurate.

[b]SOT: sensory organization test.

[c]Vis: visual.

[d]Som: somatosensory.

[e]Ves: vestibular.

Results

Data Presentation and Assessment of the Raw Data

Boxplots of the data showing the median (thick line), IQR (box edges), and 95% CI (whiskers) for each condition were created (Figure 2). Visual inspection of the data indicated symmetry in most conditions and increased variability in the more challenging conditions (conditions 4-6).

Figure 2. Boxplots of all data collected in comparable SOT (light) and VR (dark) conditions. The median value (thick line), IQR (box edges), and 95% CI (whiskers) are indicated. SOT: sensory organization test; VR: virtual reality.
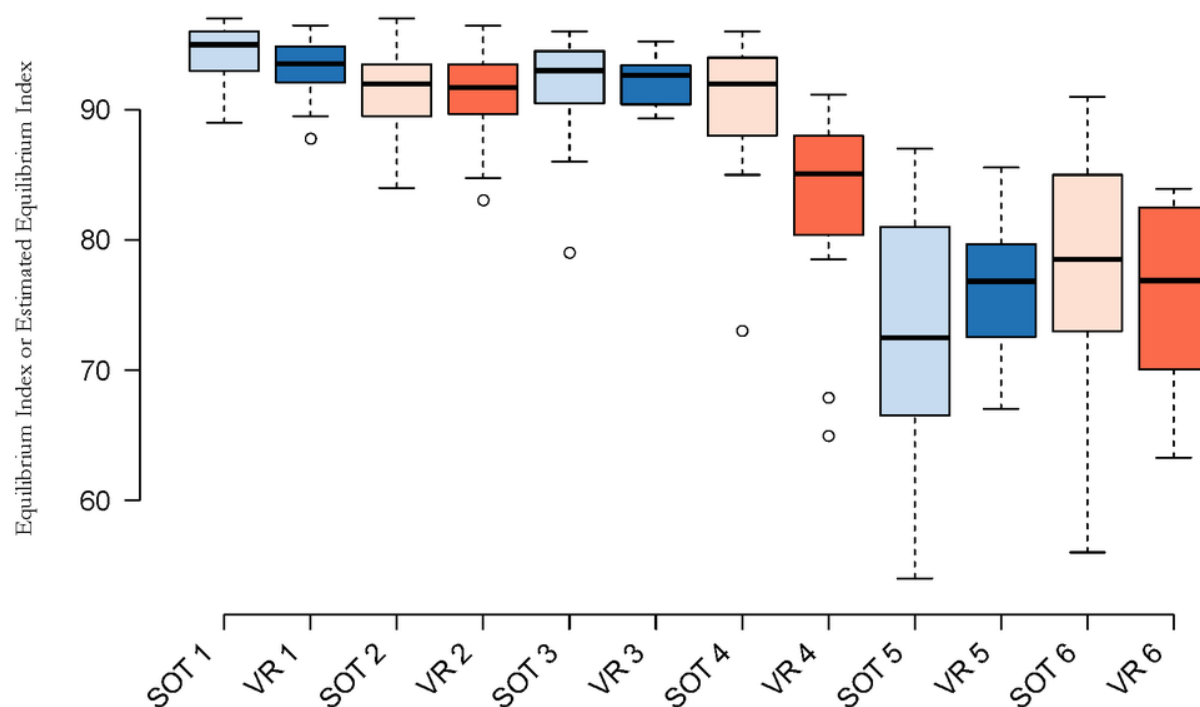
Reliability of the eEI

Intraclass correlations between EI and eEI in similar conditions were evaluated and are presented alongside Bland-Altman plots in Figure 3 [17]. SOT conditions 1, 2, 3, and 6 demonstrated moderate consistency and absolute agreement with their similar VR condition counterparts. Meanwhile, SOT conditions 4 and

5 showed poor consistency and absolute agreement with similar VR conditions. The Bland-Altman plots provide a visual representation of agreement between two measurements by plotting the absolute agreement or mean difference between measurements on the vertical axis against the average of the two measurements on the horizontal axis.

Figure 3. Bland-Altman plots comparing the equilibrium index and estimated equilibrium index from the Equitest and VR SOT, respectively. The Pearson correlation coefficient (r), intraclass correlation coefficient for absolute agreement (ICCa), and intraclass correlation coefficient for consistency (ICCc) are provided. SOT: sensory organization test; VR: virtual reality.
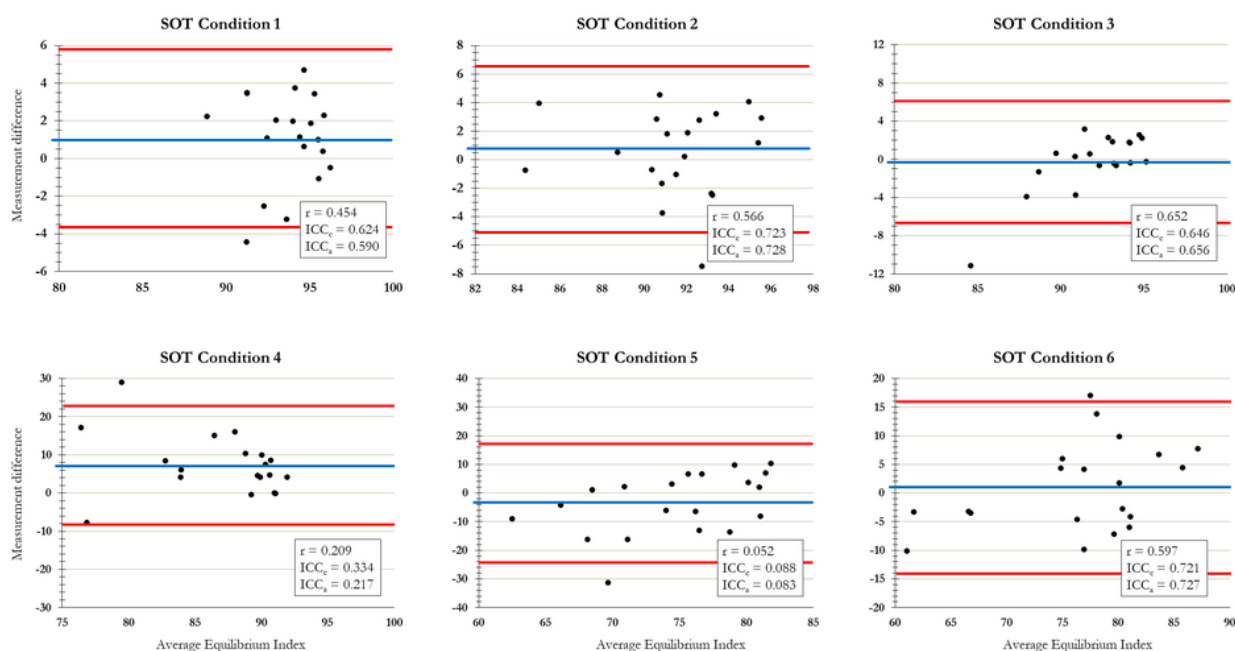
Correlation of the EI With Force Plate Measurements

Pearson correlation coefficients were calculated between the Equitest EI and balance measures derived from COP data (Table 3). Weak to moderate significant correlations were identified between EI and eEI in SOT conditions 1 (r=0.454, P=.045), 2 (r=0.566, P=.009), 3 (r=0.652, P=.002), and 6 (r=0.597, P=.005). Additionally, weak to moderate significant correlations were identified between EI and 95% ellipse area in conditions 1 (r=−0.453, P=.045), 2 (r=−0.506, P=.02), and 6 (r=−0.500, P=.03) and AP DFA α in condition 1 (r=−0.511, P=.02). No other relevant correlations were identified between the Equitest EI and balance measurements derived from the COP data.

Table 3. Pearson correlation coefficients between force plate measurements (columns) and the equilibrium index during each sensory organization test condition.

Condition	eEI[a]	95% ellipse area	Path length	AP DFA α[b]
SOT[c] 1				
r	0.454	−0.453	−0.130	−0.511
P	.045	.045	.59	.02
SOT 2				
r	0.566	−0.506	−0.400	−0.041
P	.009	.02	.08	.86
SOT 3				
r	0.652	−0.329	−0.068	−0.234
P	.002	.17	.78	.33
SOT 4				
r	0.209	−0.143	−0.332	−0.007
P	.39	.56	.16	.98
SOT 5				
r	0.052	−0.242	−0.241	0.027
P	.83	.30	.31	.91
SOT 6				
r	0.597	−0.500	−0.334	−0.174
P	.005	.03	.15	.46

[a]eEI: estimated equilibrium index.

[b]AP DFA α: anterior-posterior detrended fluctuation analysis scaling exponent alpha.

[c]SOT: sensory organization test.

Discussion

This research has demonstrated the plausibility of using VR as an alternative to the Equitest when conducting a SOT. Although not a perfect replacement, eEI demonstrated reasonable correlations and ICCs with the clinical standard in several of the SOT conditions. Continued improvements to the VR testing environment need to be made to have more confidence in its use as a potential replacement. For example, the VR device may do a good job at mimicking the visual conditions of the SOT, but the foam mat might not equivocally disrupt somatosensory information compared with the SOT. This is supported by seeing higher correlations between EI and eEI in the intact than inaccurate somatosensory conditions (conditions 1, 2, and 3 versus conditions 4 and 5). Additionally, this study identified a number of correlations between the Equitest system and typical balance measurements derived from COP data on a force plate. Aside from eEI, 95% ellipse area and AP DFA α had some correlations with the clinical test. It is not surprising that these correlations were somewhat sparse as they distinctly measure different characteristics of balance. The SOT measures only AP sway magnitude, while COP data can be used to calculate sway magnitude in the frontal and sagittal planes combined or to measure aspects of how variability is structured in an individual plane. For example, 95% ellipse area quantifies the gross postural control behavior during quiet stance [18] and AP DFA α quantifies the structure of variability within an individual's AP sway trajectory (ie, how random or deterministic the data is) [19], whereas EI evaluates how close an individual gets to a theoretical limit of stability [20]. The measures evaluated in this study were selected to represent a small array of postural control measurements, and future research should evaluate the clinical utility of individual metrics.

The recent surge in consumer-ready VR headsets has the potential to greatly reduce the cost of conducting balance assessments while also providing additional accessibility to sites outside of the clinic, for example, on the sideline during an athletic event. Likewise, using force plates opens access to

raw, processed, and derived outcome measures that take advantage of the full scope of postural dynamics and present the opportunity to have more accurate information at the clinician's disposal. In the future, it may even be possible to accurately assess balance (and gait) using only the self-contained tracking of VR headsets. This research serves as a point from which we can merge motor control assessments with the accelerating advancements in consumer technologies.

Conflicts of Interest

None declared.

References

1. Chiarovano E, Wang W, Rogers SJ, MacDougall HG, Curthoys IS, de Waele C. Balance in Virtual Reality: Effect of Age and Bilateral Vestibular Loss. Front Neurol 2017;8:5 [FREE Full text] [doi: 10.3389/fneur.2017.00005] [Medline: 28163693]

2. Robert MT, Ballaz L, Lemay M. The effect of viewing a virtual environment through a head-mounted display on balance. Gait Posture 2016 Jul;48:261-266. [doi: 10.1016/j.gaitpost.2016.06.010] [Medline: 27344394]

3. Lloréns R, Noé E, Colomer C, Alcañiz M. Effectiveness, usability, and cost-benefit of a virtual reality-based telerehabilitation program for balance recovery after stroke: a randomized controlled trial. Arch Phys Med Rehabil 2015 Mar;96(3):418-425.e2 [FREE Full text] [doi: 10.1016/j.apmr.2014.10.019] [Medline: 25448245]

4. Mirelman A, Maidan I, Deutsch JE. Virtual reality and motor imagery: promising tools for assessment and therapy in Parkinson's disease. Mov Disord 2013 Sep 15;28(11):1597-1608 [FREE Full text] [doi: 10.1002/mds.25670] [Medline: 24132848]

5. Morel M, Bideau B, Lardy J, Kulpa R. Advantages and limitations of virtual reality for balance assessment and rehabilitation. Neurophysiol Clin 2015 Nov;45(4-5):315-326 [FREE Full text] [doi: 10.1016/j.neucli.2015.09.007] [Medline: 26527045]

6. Saldana S, Marsh AP, Rejeski WJ, Haberl J, Wu P, Rosenthal S, et al. Assessing balance through the use of a low-cost head-mounted display in older adults: a pilot study. Clin Interv Aging 2017;12:1363-1370 [FREE Full text] [doi: 10.2147/CIA.S141251] [Medline: 28883717]

7. Wright WG, McDevitt J, Tierney R, Haran FJ, Appiah-Kubi KO, Dumont A. Assessing subacute mild traumatic brain injury with a portable virtual reality balance device. Disabil Rehabil 2017 Jul;39(15):1564-1572 [FREE Full text] [doi: 10.1080/09638288.2016.1226432] [Medline: 27718642]

8. de Rooij IJ, van de Port IG, Meijer JW. Effect of Virtual Reality Training on Balance and Gait Ability in Patients With Stroke: Systematic Review and Meta-Analysis. Phys Ther 2016 Dec;96(12):1905-1918 [FREE Full text] [doi: 10.2522/ptj.20160054] [Medline: 27174255]

9. Meyns P, Pans L, Plasmans K, Heyrman L, Desloovere K, Molenaers G. The Effect of Additional Virtual Reality Training on Balance in Children with Cerebral Palsy after Lower Limb Surgery: A Feasibility Study. Games Health J 2017 Feb;6(1):39-48 [FREE Full text] [doi: 10.1089/g4h.2016.0069] [Medline: 28051880]

10. Villiger M, Liviero J, Awai L, Stoop R, Pyk P, Clijsen R, et al. Home-Based Virtual Reality-Augmented Training Improves Lower Limb Muscle Strength, Balance, and Functional Mobility following Chronic Incomplete Spinal Cord Injury. Front Neurol 2017;8:635 [FREE Full text] [doi: 10.3389/fneur.2017.00635] [Medline: 29234302]

11. Yen CY, Lin KH, Hu MH, Wu RM, Lu TW, Lin CH. Effects of virtual reality-augmented balance training on sensory organization and attentional demand for postural control in people with Parkinson disease: a randomized controlled trial. Phys Ther 2011 Jun;91(6):862-874 [FREE Full text] [doi: 10.2522/ptj.20100050] [Medline: 21474638]

12. Nashner LM. EquiTest system operator's manual, version 4.04. Clackamas, OR, USA: NeuroCom International, Inc; 1992.

13. Duarte M, Freitas SM. Revision of posturography based on force plate for balance evaluation. Rev Bras Fisioter 2010;14(3):183-192. [Medline: 20730361]

14. Peng CK, Havlin S, Stanley HE, Goldberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos 1995;5(1):82-87. [doi: 10.1063/1.166141] [Medline: 11538314]

15. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012]

16. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012 Sep;24(3):69-71 [FREE Full text] [Medline: 23638278]

17. Bland J, Altman D. Measuring agreement in method comparison studies. Stat Methods Med Res 1999 Jun 01;8(2):135-160 [FREE Full text] [doi: 10.1191/096228099673819272]

18. Strang AJ, Haworth J, Hieronymus M, Walsh M, Smart LJ. Structural changes in postural sway lend insight into effects of balance training, vision, and support surface on postural control in a healthy population. Eur J Appl Physiol 2011 Jul;111(7):1485-1495 [FREE Full text] [doi: 10.1007/s00421-010-1770-6] [Medline: 21165641]

19. Rhea CK, Silver TA, Hong SL, Ryu JH, Studenka BE, Hughes CM, et al. Noise and complexity in human postural control: interpreting the different estimations of entropy. PLoS One 2011 Mar 17;6(3):e17696 [FREE Full text] [doi: 10.1371/journal.pone.0017696] [Medline: 21437281]

20. Chaudhry H, Findley T, Quigley KS, Bukiet B, Ji Z, Sims T, et al. Measures of postural stability. J Rehabil Res Dev 2004 Sep;41(5):713-720 [FREE Full text] [doi: 10.1682/jrrd.2003.09.0140] [Medline: 15558401]

Abbreviations

6DoF: six degrees of freedom
AP: anterior-posterior
COG: center of gravity
COP: center of pressure
DFA α: detrended fluctuation analysis scaling exponent alpha
eEI: estimated equilibrium index
EI: equilibrium index
ICC: intraclass correlation coefficient
SOT: sensory organization test
VR: virtual reality

<u>Original Paper</u>

Health Care Providers' Performance, Mindset, and Attitudes Toward a Neonatal Resuscitation Computer-Based Simulator: Empirical Study

Maria Cutumisu[1,2,3], PhD; Simran K Ghoman[2,4], BSc; Chang Lu[1], MEd; Siddhi D Patel[2,4], BSc; Catalina Garcia-Hidalgo[2,4], BSc; Caroline Fray[2], RN; Matthew R G Brown[2,3], PhD; Russell Greiner[3], PhD; Georg M Schmölzer[2,4], MD, PhD

[1]Department of Educational Psychology, Faculty of Education, University of Alberta, Edmonton, AB, Canada

[2]Centre for the Studies of Asphyxia and Resuscitation, Neonatal Research Unit, Royal Alexandra Hospital, Edmonton, AB, Canada

[3]Department of Computing Science, Faculty of Science, University of Alberta, Edmonton, AB, Canada

[4]Department of Pediatrics, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Maria Cutumisu, PhD
Department of Educational Psychology
Faculty of Education
University of Alberta
6-102 Education North
Edmonton, AB, T6G2G5
Canada
Phone: 1 7804925211
Email: cutumisu@ualberta.ca

Abstract

Background: Neonatal resuscitation involves a complex sequence of actions to establish an infant's cardiorespiratory function at birth. Many of these responses, which identify the best action sequence in each situation, are taught as part of the recurrent Neonatal Resuscitation Program training, but they have a low incidence in practice, which leaves health care providers (HCPs) less prepared to respond appropriately and efficiently when they do occur. Computer-based simulators are increasingly used to complement traditional training in medical education, especially in the COVID-19 pandemic era of mass transition to digital education. However, it is not known how learners' attitudes toward computer-based learning and assessment environments influence their performance.

Objective: This study explores the relation between HCPs' attitudes toward a computer-based simulator and their performance in the computer-based simulator, RETAIN (REsuscitation TrAINing), to uncover the predictors of performance in computer-based simulation environments for neonatal resuscitation.

Methods: Participants were 50 neonatal HCPs (45 females, 4 males, 1 not reported; 16 respiratory therapists, 33 registered nurses and nurse practitioners, and 1 physician) affiliated with a large university hospital. Participants completed a demographic presurvey before playing the game and an attitudinal postsurvey after completing the RETAIN game. Participants' survey responses were collected to measure attitudes toward the computer-based simulator, among other factors. Knowledge on neonatal resuscitation was assessed in each round of the game through increasingly difficult neonatal resuscitation scenarios. This study investigated the moderating role of mindset on the association between the perceived benefits of understanding the terminology used in the computer-based simulator, RETAIN, and their performance on the neonatal resuscitation tasks covered by RETAIN.

Results: The results revealed that mindset moderated the relation between participants' perceived terminology used in RETAIN and their actual performance in the game ($F_{3,44}$=4.56, R^2=0.24, adjusted R^2=0.19; P=.007; estimate=–1.19, SE=0.38, t_{44}=–3.12, 95% CI –1.96 to –0.42; P=.003). Specifically, participants who perceived the terminology useful also performed better but only when endorsing more of a growth mindset; they also performed worse when endorsing more of a fixed mindset. Most participants reported that they enjoyed playing the game. The more the HCPs agreed that the terminology in the tutorial and in the game was accessible, the better they performed in the game, but only when they reported endorsing a growth mindset exceeding the average

XSL•FO
RenderX

mindset of all the participants ($F_{3,44}=6.31$, $R^2=0.30$, adjusted $R^2=0.25$; $P=.001$; estimate$=-1.21$, SE$=0.38$, $t_{44}=-3.16$, 95% CI -1.99 to -0.44; $P=.003$).

Conclusions: Mindset moderates the strength of the relationship between HCPs' perception of the role that the terminology employed in a game simulator has on their performance and their actual performance in a computer-based simulator designed for neonatal resuscitation training. Implications of this research include the design and development of interactive learning environments that can support HCPs in performing better on neonatal resuscitation tasks.

Introduction

Background

Approximately 1 in 10 infants worldwide will require some degree of neonatal resuscitation at birth to support their circulation and breathing [1]. Infants may receive assistance across 4 categories of sequential actions: initial stabilization (provide warmth, clear airways, dry, stimulate, and reevaluate), ventilation, chest compressions, and administration of epinephrine and volume expansion [2]. These actions, including a continuous evaluation of the infant, must be performed rapidly, yet accurately. Moreover, these actions are usually performed collaboratively in teams of specialized health care providers (HCPs), which adds complexity to an already high-stress, time-sensitive task of saving an infant's life, as miscommunication can lead to errors and waste precious time [3]. Furthermore, 1% of the infants require more extensive resuscitation measures, such as chest compression and epinephrine [4].

Many HCPs may encounter these types of high-acuity low-opportunity events once in their careers. Such events require exceptional expertise, team dynamics, as well as cognitive and psychomotor acuity. Due to the rare occurrence of these highly specialized events and their collaborative, team-based nature, breakdown in HCP communication is the leading cause of neonatal death [5]. In such situations, deviation from the neonatal resuscitation protocol can occur. Specifically, it is estimated that human error causes over two-thirds of morbidity and mortality in neonatal resuscitation cases [5]. Moreover, these errors are directly proportional to the complexity of the resuscitation [6].

HCPs experience decline in skills after training over time [7]. This presents a challenge, given the limited exposure of HCPs (novice and expert alike) to challenging, realistic, complex, high-risk, and rarely occurring scenarios, compounded by infrequent hands-on experiences in traditional medical education. One potential strategy to compensate for this scarcity is periodic simulation-based training to acquire and maintain expertise [8], especially in medical education [9]. Health care education has employed simulations that represent abstractions of real-life scenarios and that incorporate expert knowledge models to provide alternative opportunities for training and learning [10]. Moreover, neonatal resuscitation guidelines recommend the use of simulation-based medical education as

a solution to mitigate the loss of skills over time and to decrease human error in the delivery room [11]. Several researchers have used computer-based simulations successfully in pediatric and neonatal resuscitation for the last 2 decades to improve performance and learning [12-16].

Simulations have several benefits over more traditional training experiences. They offer deliberate practice and experiential learning opportunities associated with better learning, especially when they foster a safe and supportive environment, where making mistakes is welcomed as an opportunity to uncover and remedy gaps in knowledge [17,18]. Simulation is a safer alternative, as no real patients are required. Furthermore, it puts neither the patient nor the trainee at risk, especially in neonatal resuscitation when the outcome in the real situation may be the loss of life or in simulating infectious diseases when the outcome may be the contamination of the trainee. In neonatal resuscitation, simulations showed several benefits for HCPs [19-21]. In some cases, simulations can be used anytime and anywhere and can reproduce rarely occurring training scenarios and can tune the difficulty and complexity of the scenario to exemplify the phenomena of interest. Simulations can also provide immediate expert feedback and diagnostic assessment, which were also found to support learning. Importantly, the experiential nature of the simulations and the similarity of simulated scenarios with real-life situations may help the participant transfer the skills learned, for instance, from the simulation to the delivery room [22]. However, very few studies were conducted to specifically target transfer. For instance, in a recent scoping review of medical student training in eHealth, none of the articles reviewed aimed to demonstrate that the eHealth training of medical students transferred outside the simulation environment [23].

The REsuscitation TrAINing Simulator

The RETAIN (REsuscitation TrAINing) simulator employed in this study is a computer role-playing game [24,25] that was designed to support novice HCPs in acquiring neonatal resuscitation knowledge and to assist expert HCPs in refreshing their knowledge, especially by exposing them to novel, complex, and rarely occurring scenarios, in between taking the Neonatal Resuscitation Program (NRP) refresher courses [26]. The game is also relatively short, up to 10 minutes, which fits in an HCP's busy schedule, and it is easily accessible, as it only requires a computer.

Although there is a paucity of computer-based simulators, they have been found to improve knowledge and decision-making skills [27-30]. However, few studies have examined participants' experiences in these environments and the attitudes (eg, mindsets or theories of intelligence) they bring to these tasks [31]. Thus, this study adds to the research body on neonatal resuscitation simulators by analyzing the survey responses and computer-based game simulator performance of HCPs to gain an insight into their perceptions of the simulation environment, their performance on these tasks, and the relationship between their attitudes and performance. This exploration is prompted by the belief that individuals who endorse a growth mindset are more interested in mastery and work harder to overcome barriers and achieve their goals, as they believe that intelligence and ability are malleable; concomitantly, those who endorse a fixed mindset are more interested in performance and do not work as hard, as they do not believe that they can change their abilities [31]. The objectives of this study were (1) to determine whether attitudes toward the simulator hinder or enhance HCPs' performance in a neonatal resuscitation computer-based game simulation and (2) to examine whether the HCP's mindset moderates this relationship between attitudes and performance. We hypothesized that HCPs' mindset would strengthen the relation between their perceived performance, given their understanding of the terminology used in the neonatal resuscitation simulator and their actual performance in this environment.

This study contributes to understanding the influence of HCPs' mindset and perceptions of computer-based game simulations for neonatal resuscitation on their performance. Moreover, these noncognitive factors are examined in conjunction with HCPs' performance in increasingly difficult neonatal resuscitation scenarios, providing an insight into cognitive factors and into the impact of attitudes on performance. Lessons learned from this study may help medical education instructors incorporate computer-based game simulations in their training and instructional practice.

Methods

Participants

Fifty HCPs (45 females, 4 males, and 1 not reported), who had completed their NRP [26] training within the 24 months preceding this study, were recruited voluntarily from the neonatal intensive care unit (NICU) at the Royal Alexandra Hospital, Edmonton, Canada, a tertiary perinatal center delivering over 6500 infants every year. The sample consisted of 16 respiratory therapists, 33 registered nurses and nurse practitioners, and 1 physician, which was representative of the HCP population within the NICU. The study was performed at the bedside in the NICU and it was approved by the Human Research Ethics Board at the University of Alberta (Pro00064234). Written informed consent was obtained from all HCPs prior to participation and no participants were excluded. None of the participants had any prior experience with the RETAIN computer-game simulation.

Study Setup

The study was conducted based on a computer-based game simulator RETAIN (for HCPs) at the simulation laboratory at the Centre for the Studies of Asphyxia and Resuscitation [32]. RETAIN was designed to support HCPs in practicing their neonatal resuscitation knowledge in between regular NRP [26] refresher courses. In the RETAIN computer-based game simulator, participants, assuming the role of a neonatologist, tackled increasingly difficult simulated neonatal resuscitation scenarios in each of the 3 rounds of the game, each of which required first repeating and then extending the steps taken on the previous rounds. Their avatar (ie, a medical doctor) needed to perform 4 categories of sequential actions: initial stabilization, ventilation, chest compressions, and administration of epinephrine and volume expansion to assist an infant at birth. For example, the last game round required the player to perform mask ventilation, chest compression, and administer epinephrine to stabilize the infant. The participants had a limited amount of time to complete the neonatal resuscitation scenario presented in each game round, as the game provided a countdown that simulated the urgency of a real-world, high-stakes, fast-paced delivery room. Players were allowed to advance to the next game round only if they made at most 4 mistakes in the current game round. Otherwise, they were required to repeat that round.

Procedure and Data Collection

All participants completed 2 surveys—1 before and 1 after completing the computer-based simulation. The presurvey included demographic and educational background items (eg, time in months since the participant's last NRP course), whereas the postsurvey consisted of attitudinal items, including their views on the current computer simulation and mindset (eg, the terminology or phrasing used did not impede your ability to complete the steps). Attitudinal items included in the postsurvey used a 5-point Likert scale (1: Strongly disagree, 2: Disagree, 3: Neutral, 4: Agree, and 5: Strongly agree).

Between the surveys, the participants played the RETAIN computer game. This particular game version was implemented using the Aurora game engine of the award-winning *Neverwinter Nights* computer role-playing game [33]. The game started with a short tutorial presenting a practice scenario that familiarized players with the mechanics of the game. Then, the participants played 3 consecutive game rounds that presented resuscitation scenarios of increasing difficulty, each encompassing the steps taken in the previous rounds. The game took an average of 8.47 minutes to play. More details about this game are presented in previous studies [24,25,34]. Learning analytics regarding participants' performance and behaviors were tracked within the computer-game simulator.

Performance Measures

Number of Tries: The outcome variable employed in this study represents the number of tries performed in the game. This measure ranged from 32 (ie, the participant solved all the scenarios from the first try) to 54 (ie, the participant took more tries to solve the game scenarios).

Attitudinal Measures

All the following items were rated on a Likert scale from 1 (*Strongly disagree*) to 5 (*Strongly agree*), as mentioned above.

1. Enjoyment: This variable captured the participants' response values to the following statement, that is, "Did you enjoy playing this game?"
2. Tutorial Terminology: The predictor variable captured the participants' response values to the following statement, that is, "The terminology used did not impede your ability to complete the steps."
3. Terminology Used: We also considered another predictor variable that captured the participants' response values to the following statement about the entire game, that is, "The terminology or phrasing used did not impede your ability to complete the steps."
4. Mindset: The moderator variable captured the participants' response values to 4 items that probed participants' theories of intelligence (ie, mindsets) [31]. Two items were related to fixed mindset or the belief that intelligence is fixed (Fixed Mindset 1: "You can't really do much to change how good you are at your job" and Fixed Mindset 2: "You can learn new things, but you cannot really change how good you are at your job"), while 2 other items were related to growth mindset or the belief that intelligence is malleable (Growth Mindset 1: "You can always change how good you are at your job" and Growth Mindset 2: "You can get better at your job with practice"). All items were positively stated, except for the 2 fixed-mindset items, which were reverse-coded. Then, a Mindset variable was obtained by adding the growth-mindset items and the reversed fixed-mindset items: 12 − (Fixed Mindset 1 + Fixed Mindset 2) + (Growth Mindset 1 + Growth Mindset 2).

Statistical Analyses

All analyses were performed using version 4.0.2 of the R [35] statistical software. They included descriptive statistics and tests of association, including correlation and regression analyses, to test the moderation effect for continuous interactions.

Descriptive analyses: First, we computed the summary statistics of the study variables by using the *summary* function and the *describe* function of the *Hmisc* [36] package in R. Second, we explored the relationships between the variables included in this study. Specifically, we conducted two-fold correlation analyses between the outcome variable and the 2 mean-centered predictor and moderator variables: (1) Spearman correlations using the *corr.test* function of the *psych* [37] package, which also generates confidence intervals, and (2) robust correlations to compute the percentage-bend correlation coefficient, using the *pbcor* function of the *WRS2* [38] package in R.

Multiple regression analyses: These analyses, which were conducted using the *lm* built-in function in R [35], probed whether the moderator influenced the strength of the relation between each predictor variable and the outcome variables employed in this study. First, linearity assumptions of the multiple linear regression analysis were tested using the *gvlma* [39] package in R. Multicollinearity tests were conducted to ascertain whether multicollinearity was problematic for any of the models. The updated quantile-comparison plots for the robust models are shown in Figures S2, S3, S5, and S6 in Multimedia Appendix 1. Then, the Johnson-Neyman [40,41] technique was conducted to test the interaction, as the variables involved were continuous. We used the *sim_slopes* and the *johnson_neyman* functions of the *interactions* [42] (former *jtools*) R package to determine the *regions of significance* (ie, the precise range of the moderator values for which the main effect of the predictor on the outcome is statistically significantly different from zero) for simple slopes. The Johnson-Neyman technique displays 95% confidence bands around the regression line (ie, representing a multiplicative interaction effect model) showing how the main effect varies across the values of a moderator. Both the predictor and the moderator variables were mean-centered for this analysis to account for the multicollinearity caused by the association between the independent variables and for ease of result interpretation.

Results

Descriptive Analyses

Table 1 presents the descriptive statistics together with the correlation results for both Spearman and robust correlations.

Table 1. Descriptive statistics for the study variables, Spearman correlation coefficients and their confidence intervals for the study variables, and robust correlation coefficients.

Variable	Tutorial Terminology	Terminology Used	Mindset	Number of Tries
Spearman correlations				
Tutorial Terminology	—[a]			
Terminology Used	0.75[b] [0.60, 0.85]	—		
Mindset	0.20 [–0.08, 0.45]	0.12 [–0.17, 0.38]	—	
Number of Tries	0.08 [–0.20, 0.35]	0.09 [–0.20, 0.36]	–0.15 [–0.41, 0.14]	—
Robust correlations				
Tutorial Terminology	—			
Terminology Used	0.73[b]	—		
Mindset	0.23	0.14	—	
Number of Tries	0.11	0.09	–0.15	—
Participants (n)	50	50	48	50
Mean (SD)	4.46 (0.61)	4.30 (0.71)	17.79 (2.32)	35.72 (4.16)
Minimum	2	2	12	32
Maximum	5	5	20	54
Skewness	–1.15	–1.15	–0.86	2.16
Kurtosis	2.64	2.11	–0.01	5.69
SE	0.09	0.10	0.34	0.59

[a]Not applicable.

[b]The correlation was significant at $P<.001$.

Figure 1 shows that most participants agreed or strongly agreed with the statements regarding *Tutorial Terminology* and *Growth Mindset* but disagreed and strongly disagreed with the statements related to *Fixed Mindset*. Additionally, most participants reported that they enjoyed the game, with a mean (SD) of 4.04 (0.68) as well as a median and a mode of 4 on a self-reported item measuring participants' gameplay enjoyment.

Figure 1. Responses on a 5-point Likert scale to the items included in this study. 1: Strongly disagree; 2: Disagree; 3: Neutral; 4: Agree; and 5: Strongly agree.
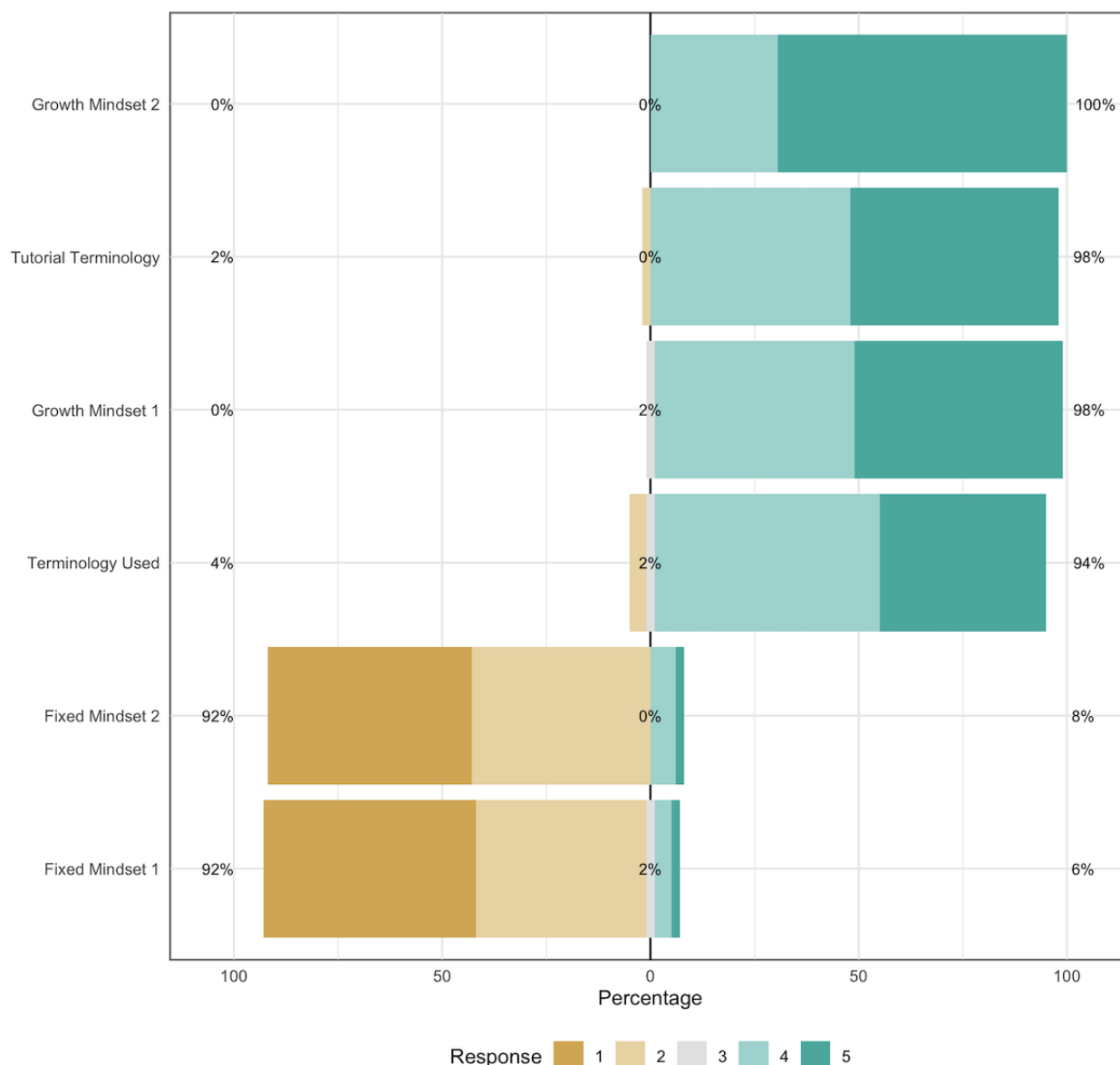
Multiple Regression Analyses

We conducted a moderation analysis to examine whether the relationship between the continuous predictor, *Tutorial Terminology* in Model 1 (*Terminology Used* in Model 2, respectively) and the dependent continuous variable, *Number of Tries*, was moderated by the continuous variable, *Mindset*.

Multicollinearity tests yielded variance inflation factor (VIF) values near 1.0 and less than 5.0 for both Model 1 (VIFpredictor=1.14, VIFmoderator=1.01, and VIFinteraction=1.13) and Model 2 (VIFpredictor=1.29, VIFmoderator=1.00, and VIFinteraction=1.29), indicating that multicollinearity was not problematic for any of the 2 models.

The Shapiro-Wilk normality test revealed that the outcome variable was not normally distributed (W=0.75, *P*<.001). The residuals of Model 1 and Model 2 were also not normally

distributed, as shown in the quantile-comparison plot of Figure S1 and Figure S4, respectively, included in Multimedia Appendix 1. Thus, a robust linear regression, which also provides robustness to outliers, was conducted using the *lmrob* function of the *robustbase* [43] packages in R for heteroscedasticity-robust fitting of linear regression models to compute a fast M estimator [44,45].

Model 1: Tutorial Terminology, Mindset, and Number of Tries

The findings shown in Table 2 revealed that the more the HCPs agreed that the *Tutorial Terminology* did not impede their ability to complete the tutorial scenarios, the fewer tries they needed to complete the game, but this association was significant only for those who endorsed more of a growth mindset (ie, exceeded the average mindset value across the sample).

Table 2. Johnson-Neyman moderator analysis for model 1: Mindset moderates the relationship between Tutorial Terminology and Number of Tries across the game[a].

Effect (n=48)	Estimate	SE	t (df)	95% CI	P value
Coefficients					
Intercept	35.67	0.50	70.91 (44)	[34.65, 36.68]	*<.001* [b]
Tutorial Terminology	–1.36	0.87	–1.55 (44)	[–3.12, 0.40]	.13
Mindset	–0.16	0.22	–0.75 (44)	[–0.60, 0.28]	.46
Tutorial Terminology:Mindset	–1.21	0.38	–3.16 (44)	[–1.99, –0.44]	.003
Simple slopes analysis					
When moderator[c] = –2.32 (-1SD)					
Slope of moderator	1.46	1.44	1.01 (44)	N/A[d]	.32
Conditional intercept	36.01	0.71	50.56 (44)	N/A	*<.001*
When moderator = 0.00 (Mean)					
Slope of moderator	–1.36	0.87	–1.55 (44)	N/A	.13
Conditional intercept	35.70	0.50	71.11 (44)	N/A	*<.001*
When moderator = 2.32 (+1SD)					
Slope of moderator	–4.18	1.02	–4.11 (44)	N/A	*<.001*
Conditional intercept	35.38	0.72	49.35 (44)	N/A	*<.001*
Robust linear regression					
Intercept	34.50	0.39	88.09 (44)	[33.72, 35.29]	*<.001*
Tutorial Terminology	1.31	0.68	1.92 (44)	[–0.07, 2.69]	.06
Mindset	–0.13	0.13	–0.96 (44)	[–0.39, 0.14]	.34
Tutorial Terminology:Mindset	–0.72	0.27	–2.70 (44)	[–1.26, –0.18]	.01

[a]Model fit: $F_{3,44}=6.31$, $R^2=0.30$, adjusted $R^2=0.25$; $P=.001$.

[b]The values in italics were significant at $P<.001$.

[c]The moderator was the centered mindset variable.

[d]N/A: not applicable.

The Johnson-Neyman interval indicated that when *Mindset* was outside the interval [–5.10, 0.28], the slope of *Tutorial Terminology* was statistically significant at the *P*<.05 level, with the range of observed values of *Mindset* being [–5.79, 2.21]. Specifically, this relationship was significant when the value of the mean-centered *Mindset* variable was included in the intervals [–5.79, –5.10) or (0.28, 2.21], which corresponds to the value of the noncentered *Mindset* variable being included in the intervals [12, 12.69) or (18.07, 20], as its mean was 17.79 as shown in Table 1. The linear model fitted using an MM-type estimator (ie, the M estimator) yielded the same results as the original nonrobust linear regression model and it identified 4 outlier observations, as shown in Table 2.

Figure 2 displays a plot of conditional effects showing how the effect of the influence exerted by the independent variable (*Tutorial Terminology*) on the dependent variable (*Number of Tries*) is conditional on the full range of the moderator (*Mindset*). This figure displays the adjusted effect of the *Tutorial Terminology* on *Number of Tries* (y axis) across all continuous values of *Mindset* (x axis). For any values of the moderator for which the confidence bands do not contain 0, the effect of the predictor on the outcome is significantly different from 0 at the *P*=.05 level.

Figure 3 shows the effect of *Tutorial Terminology* (x axis) on the *Number of Tries* (y axis) at 3 levels of *Mindset*: low (1 standard deviation lower than the mean), moderate (mean), and high (1 standard deviation higher than the mean).

Figure 2. The Johnson-Neyman interaction intervals for Model 1 with Tutorial Terminology as a predictor. The region of significance is determined by the locations where the upper and lower bounds of the 95% confidence interval intersect zero.

Figure 3. Model 1 interaction plot: the relationship between the Tutorial Terminology predictor and the Number of Tries dependent variable is significant only for very low or very high levels of the Mindset moderator.
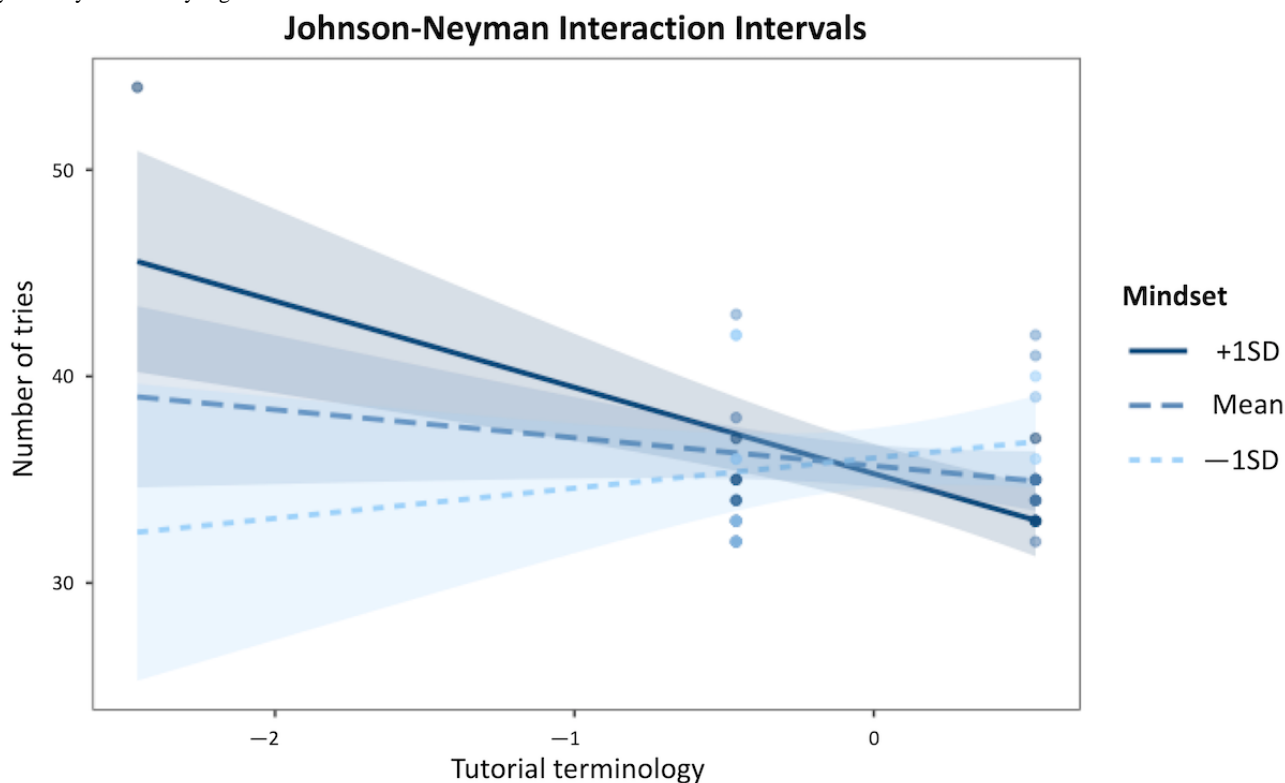
Model 2: Terminology Used, Mindset, and Number of Tries

The findings shown in Table 3 also revealed that the more the HCPs agreed that the overall *Terminology Used* in the game did not impede their ability to complete the steps, the fewer tries they took to complete the game, but this association was significant only for those who endorsed more of a growth mindset (ie, exceeded the average mindset value).

Table 3. Johnson-Neyman moderator analysis for model 2: Mindset moderates the relationship between Terminology Used and Number of Tries across the game[a].

Effect (n=48)	Estimate	SE	t (df)	95% CI	P value
Coefficients					
Intercept	35.56	0.52	68.05 (44)	[34.51, 36.61]	*<.001* [b]
Terminology Used	–0.01	0.84	–1.01 (44)	[–1.70, 1.69]	.99
Mindset	–0.21	0.23	–0.92 (44)	[–0.67, 0.25]	.36
Terminology Used:Mindset	–1.19	0.38	–3.12 (44)	[–1.96, –0.42]	.003
Simple slopes analysis					
When moderator[c] = –2.32 (–1SD)					
Slope of moderator	2.77	1.48	1.86 (44)	N/A[d]	.07
Conditional intercept	36.02	0.74	48.45 (44)	N/A	*<.001*
When moderator = 0.00 (Mean)					
Slope of moderator	–0.01	0.84	–0.01 (44)	N/A	.99
Conditional intercept	35.56	0.52	68.06 (44)	N/A	*<.001*
When moderator = 2.32 (+1SD)					
Slope of moderator	–2.78	0.89	–3.13 (44)	N/A	*<.001*
Conditional intercept	35.10	0.74	47.21 (44)	N/A	*<.001*
Robust linear regression					
Intercept	34.72	0.53	66.12 (44)	[33.66, 35.78]	*<.001*
Terminology Used	1.40	0.72	1.95 (44)	[–0.04, 2.84]	.06
Mindset	–0.28	0.18	–1.60 (44)	[–0.64, 0.07]	.12
Terminology Used:Mindset	–0.66	0.29	–2.30 (44)	[–1.23, –0.08]	*.03*

[a]Model fit: $F_{3,44}=4.56$, $R^2=0.24$, adjusted $R^2=0.19$; $P=.007$.

[b]The values in italics were significant at $P<.001$.

[c]The moderator was the centered mindset variable.

[d]N/A: not applicable.

The Johnson-Neyman interval yielded that when *Mindset* was outside the interval [–2.76, 1.25], the slope of *Terminology Used* was statistically significant at the $P<.05$ level, with the range of observed values of *Mindset* being [–5.79, 2.21], as in Model 1. Specifically, this relationship was significant when the value of the mean-centered *Mindset* variable was included in the intervals [–5.79, –2.76) or (1.25, 2.21], which is equivalent to the value of the noncentered *Mindset* variable being included in the intervals [12, 15.03) or (19.04, 20], as its mean was 17.79 as shown in Table 1. As before, the linear model fitted using an MM-estimator yielded the same results as the original nonrobust linear regression model and it identified 1 outlier, as shown in Table 3. Figure 4 shows a plot of conditional effects for Model 2. Figure 5 shows the effect of *Terminology Used* (x axis) on *Number of Tries* (y axis) for 3 levels of *Mindset*.

Figure 4. The Johnson-Neyman interaction intervals for Model 2 with Terminology Used as a predictor. The narrow area of the confidence bands represents smaller errors of estimate.
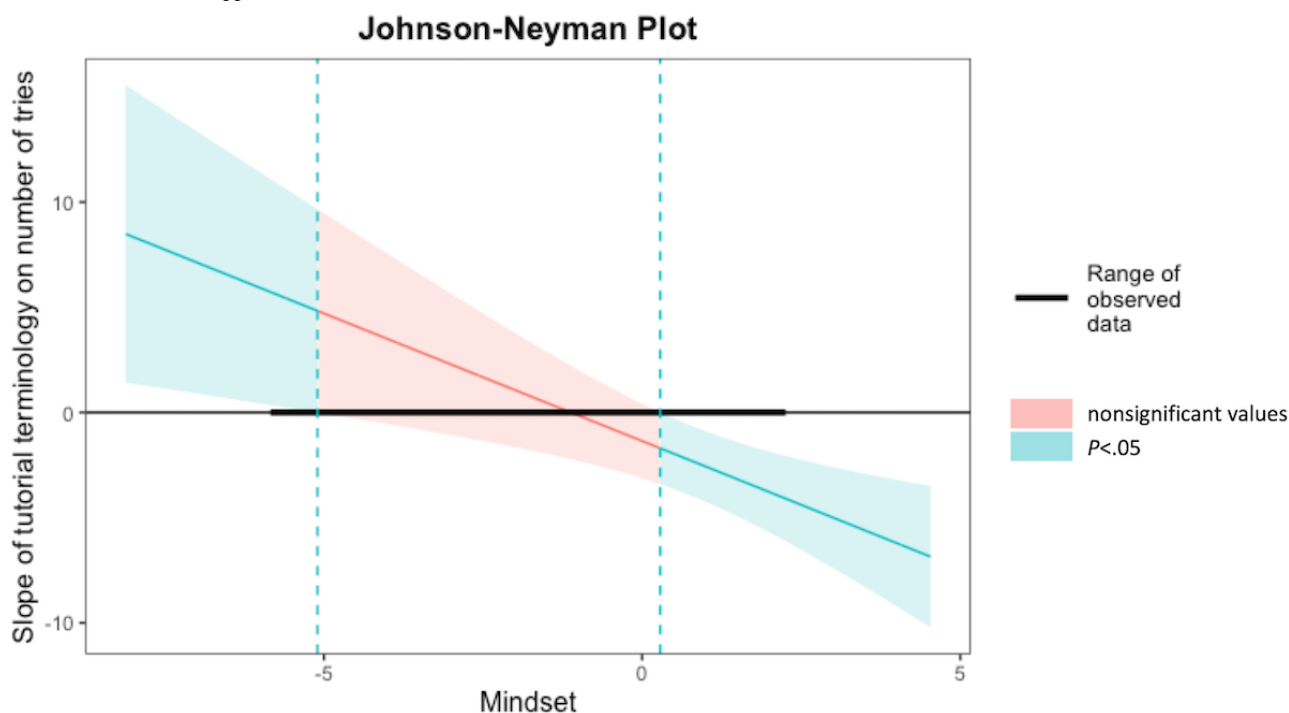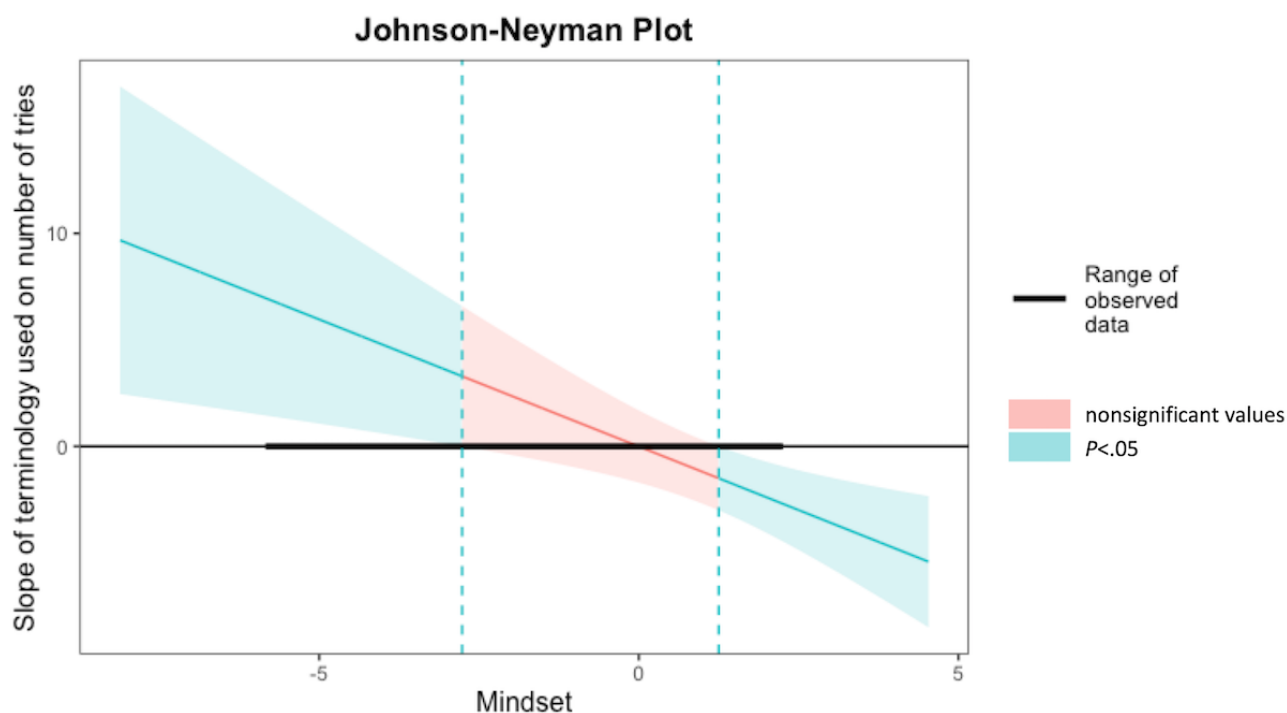

Figure 5. Model 2 interaction plot: the relationship between the Terminology Used predictor and the Number of Tries dependent variable is significant only for very low or very high levels of the Mindset moderator.
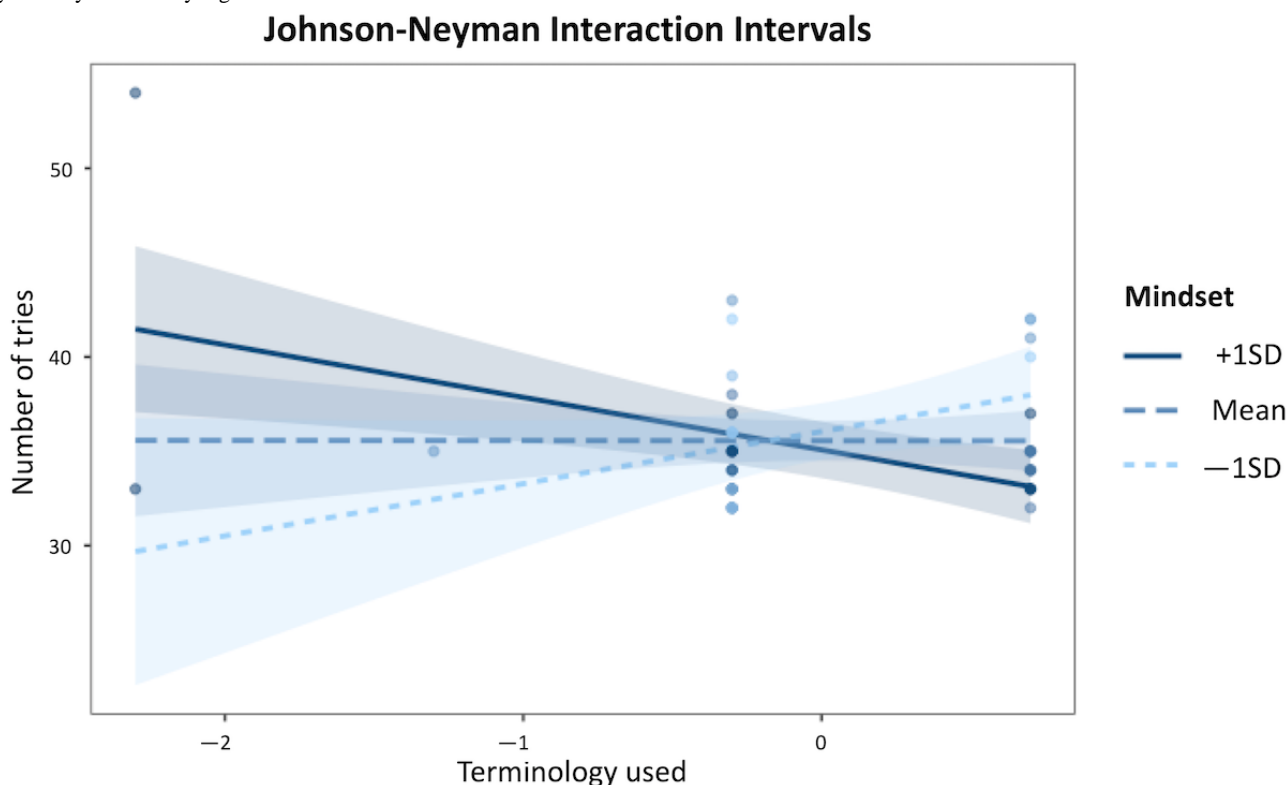
Discussion

Primary Outcomes

The results of this study revealed that the more the HCPs agreed that the terminology used in both the tutorial and the game did not impede their game experience, the better they performed in the game, but this was only when they endorsed growth-mindset levels exceeding the average growth mindset. Although there have been recent advances in computer-based simulations in terms of their high fidelity and realistic portrayal of real-world environments, few studies have explored HCPs' attitudes toward computer-based simulators, and even fewer have linked their attitudes to mindset and performance. The findings of this study

reveal that most participants agreed or strongly agreed that the terminology used in the tutorial and in the game did not impede their ability to successfully complete the game. Moreover, the results show that participants strongly endorsed growth mindset and weakly endorsed fixed mindset. Further, the findings show that HCPs largely enjoyed playing the game, which was also echoed in other computer-based simulations for neonatal resuscitation [46,47]. Together with the results regarding the use of terminology in the game, the findings suggest that HCPs appreciate the guided apprenticeship and realism of the simulator. This result is echoed in the literature, with several studies in anesthesiology showing that the use of simulators enhanced participants' understanding [48].

The results of this study also show that the more the participants agreed with statements about the terminology used in both the tutorial and in the game, the fewer tries they needed to complete the game (ie, the better they performed in the game) but only when they endorsed more of a growth mindset. This is represented by the right region of significance in the Johnson-Neyman interaction plots and by the negative slope when mindset is around 1 standard deviation above the mean. Conversely, the more the participants agreed with statements about the terminology used in both the tutorial and in the game, the more tries they made in the game (ie, the worse they performed in the game) but only when they endorsed more of a fixed mindset. This is represented by the left region of significance in the Johnson-Neyman interaction plots and by the positive slope when mindset is around 1 standard deviation below the mean. A previous research study found that endorsing a growth mindset moderated the relationship between the time elapsed from the last NRP refresher course and the number of mistakes in a neonatal resuscitation task performed in a video game [25], thereby showing the mindset is an important factor to consider in relation to performance. However, that study used 2 constructs for mindset (1 for growth and 1 for fixed mindset), in contrast to 1 mindset continuum employed in this study. Moreover, recent research studies suggest that mindset may influence performance differentially if it is generic or domain-specific [49].

Computer-based game simulators for neonatal resuscitation such as RETAIN provide many opportunities for HCPs to acquire knowledge and practice their skills in a low-cost, low-stakes, and enjoyable learning environment that provides guidance and feedback, opportunities for repetition of missteps, and various levels of difficulty. They may help HCPs interpret complex situations better and thus complement routine refresher courses. This study adds to the literature that shows support for the integration of computer-based simulations in medical education by examining both noncognitive and cognitive factors, as well as the relationship between them.

Future work could also explore collaborative computer-based simulators of neonatal resuscitation. This study used a single-player game simulator, but extensions to multiplayer experiences could test the hypothesis that miscommunications among the team members decrease overall individual performance and infant outcomes.

Conclusions

This study examined 50 HCPs' performance as well as their mindsets and attitudes toward the terminology employed in the RETAIN computer-based simulator. The results revealed that the more the HCPs agreed that the terminology used in the tutorial and in the game did not impede their game experience, the better they performed in the game, but only when they endorsed more of a growth mindset. Conversely, the more they agreed with the terminology statements, the worse they performed in the game, but only when they endorsed more of a fixed mindset. These findings suggest that HCPs' noncognitive factors such as the perception of the game and their mindsets have a significant impact on their actual performance in the game medium. Thus, one research direction could explore complex cognitive and noncognitive processes that may drive a set of HCP behaviors leading to better neonatal resuscitation performance. As computer-based game simulators are less costly and more practical than traditional training in neonatal resuscitation, future research will examine the role of computer-based game simulators in improving the safety of neonatal resuscitation procedures in the delivery room by investigating potential knowledge transfer, especially for high-risk deliveries.

Conflicts of Interest

GMS and MRGB have registered the RETAIN table-top simulator (Tech ID 2017083) and the RETAIN digital simulator under Canadian copyright (Tech ID-2017086). All other authors declare no conflicts.

Multimedia Appendix 1

Supplementary figures.
[DOCX File , 693 KB - games_v8i4e21855_app1.docx]

References

1. Aziz K, Chadwick M, Baker M, Andrews W. Ante- and intra-partum factors that predict increased need for neonatal resuscitation. Resuscitation 2008 Dec;79(3):444-452. [doi: 10.1016/j.resuscitation.2008.08.004] [Medline: 18952348]

2. Kattwinkel J, Perlman JM, Aziz K, Colby C, Fairchild K, Gallagher J, et al. Part 15: neonatal resuscitation: 2010 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2010 Nov 02;122(18 Suppl 3):S909-S919. [doi: 10.1161/CIRCULATIONAHA.110.971119] [Medline: 20956231]

3. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. Acta Anaesthesiol Scand 2009 Feb;53(2):143-151. [doi: 10.1111/j.1399-6576.2008.01717.x] [Medline: 19032571]

4. Barber CA, Wyckoff MH. Use and efficacy of endotracheal versus intravenous epinephrine during neonatal cardiopulmonary resuscitation in the delivery room. Pediatrics 2006 Sep;118(3):1028-1034. [doi: 10.1542/peds.2006-0416] [Medline: 16950994]

5. No authors listed. Sentinel event alert issue 30--July 21, 2004. Preventing infant death and injury during delivery. Adv Neonatal Care 2004 Aug;4(4):180-181. [Medline: 15368207]

6. Yamada NK, Yaeger KA, Halamek LP. Analysis and classification of errors made by teams during neonatal resuscitation. Resuscitation 2015 Nov;96:109-113. [doi: 10.1016/j.resuscitation.2015.07.048] [Medline: 26282500]

7. Cusack J, Fawke J. Neonatal resuscitation: are your trainees performing as you think they are? A retrospective review of a structured resuscitation assessment for neonatal medical trainees over an 8-year period. Arch Dis Child Fetal Neonatal Ed 2012 Jul;97(4):F246-F248. [doi: 10.1136/archdischild-2011-300548] [Medline: 22102634]

8. Hoffman RR, Ward P, Feltovich PJ, DiBello L, Fiore SM, Andrews DH. Accelerated Expertise: Training for High Proficiency in a Complex World. Abingdon, UK: Psychology Press; 2014:1-272.

9. Streufert S, Satish U, Barach P. Improving Medical Care: The Use of Simulation Technology. Simulation & Gaming 2016 Aug 18;32(2):164-174. [doi: 10.1177/104687810103200205]

10. Patterson MD, Militello LG, Bunger A, Taylor RG, Wheeler DS, Klein G, et al. Leveraging the Critical Decision Method to Develop Simulation-Based Training for Early Recognition of Sepsis. Journal of Cognitive Engineering and Decision Making 2016 Feb 29;10(1):36-56. [doi: 10.1177/1555343416629520]

11. Perlman JM, Wyllie J, Kattwinkel J, Wyckoff MH, Aziz K, Guinsburg R, Neonatal Resuscitation Chapter Collaborators. Part 7: Neonatal Resuscitation: 2015 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science With Treatment Recommendations. Circulation 2015 Oct 20;132(16 Suppl 1):S204-S241. [doi: 10.1161/CIR.0000000000000276] [Medline: 26472855]

12. Biese KJ, Moro-Sutherland D, Furberg RD, Downing B, Glickman L, Murphy A, et al. Using screen-based simulation to improve performance during pediatric resuscitation. Acad Emerg Med 2009 Dec;16 Suppl 2:S71-S75 [FREE Full text] [doi: 10.1111/j.1553-2712.2009.00590.x] [Medline: 20053216]

13. Bracq M, Michinov E, Jannin P. Virtual Reality Simulation in Nontechnical Skills Training for Healthcare Professionals: A Systematic Review. Simul Healthc 2019 Jun;14(3):188-194. [doi: 10.1097/SIH.0000000000000347] [Medline: 30601464]

14. Chang TP, Weiner D. Screen-Based Simulation and Virtual Reality for Pediatric Emergency Medicine. Clinical Pediatric Emergency Medicine 2016 Sep;17(3):224-230. [doi: 10.1016/j.cpem.2016.05.002]

15. Persson J. A review of the design and development processes of simulation for training in healthcare - A technology-centered versus a human-centered perspective. Appl Ergon 2017 Jan;58:314-326. [doi: 10.1016/j.apergo.2016.07.007] [Medline: 27633228]

16. Sawyer T, Umoren RA, Gray MM. Neonatal resuscitation: advances in training and practice. Adv Med Educ Pract 2017;8:11-19 [FREE Full text] [doi: 10.2147/AMEP.S109099] [Medline: 28096704]

17. Lewis R, Strachan A, Smith MM. Is high fidelity simulation the most effective method for the development of non-technical skills in nursing? A review of the current evidence. Open Nurs J 2012;6:82-89 [FREE Full text] [doi: 10.2174/1874434601206010082] [Medline: 22893783]

18. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010 Jan;44(1):50-63. [doi: 10.1111/j.1365-2923.2009.03547.x] [Medline: 20078756]

19. Bruno CJ, Angert R, Rosen O, Lee C, Vega M, Kim M, et al. Simulation as a tool for improving acquisition of neonatal resuscitation skills for obstetric residents. J Matern Fetal Neonatal Med 2016;29(16):2625-2629. [doi: 10.3109/14767058.2015.1095179] [Medline: 26456348]

20. Lopreiato JO, Sawyer T. Simulation-based medical education in pediatrics. Acad Pediatr 2015;15(2):134-142. [doi: 10.1016/j.acap.2014.10.010] [Medline: 25748973]

21. Malmström B, Nohlert E, Ewald U, Widarsson M. Simulation-based team training improved the self-assessed ability of physicians, nurses and midwives to perform neonatal resuscitation. Acta Paediatr 2017 Aug;106(8):1273-1279. [doi: 10.1111/apa.13861] [Medline: 28370414]

22.   Weaver SJ, Salas E, Lyons R, Lazzara EH, Rosen MA, Diazgranados D, et al. Simulation-based team training at the sharp
       end: A qualitative study of simulation-based team training design, implementation, and evaluation in healthcare. J Emerg
       Trauma Shock 2010 Oct;3(4):369-377 [FREE Full text] [doi: 10.4103/0974-2700.70754] [Medline: 21063560]
23.   Echelard J, Méthot F, Nguyen H, Pomey M. Medical Student Training in eHealth: Scoping Review. JMIR Med Educ 2020
       Sep 11;6(2):e20027 [FREE Full text] [doi: 10.2196/20027] [Medline: 32915154]
24.   Bulitko V, Hong J, Kumaran K, et al.. RETAIN: A neonatal resuscitation trainer built in an undergraduate video-game
       class. 2015 Jul 02. URL: https://arxiv.org/pdf/1507.00956.pdf [accessed 2020-11-28]
25.   Cutumisu M, Brown MRG, Fray C, Schmölzer GM. Growth Mindset Moderates the Effect of the Neonatal Resuscitation
       Program on Performance in a Computer-Based Game Training Simulation. Front Pediatr 2018;6:195 [FREE Full text] [doi:
       10.3389/fped.2018.00195] [Medline: 30023355]
26.   Weiner GM, Zaichkin J. Textbook of neonatal resuscitation (NRP). 7th ed. 2016. URL: http://gynecology.sbmu.ac.ir/
       uploads/4_5922741106407112942.pdf [accessed 2020-03-01]
27.   Schmölzer GM, Kumar M, Pichler G, Aziz K, O'Reilly M, Cheung P. Non-invasive versus invasive respiratory support in
       preterm infants at birth: systematic review and meta-analysis. BMJ 2013 Oct 17;347:f5980 [FREE Full text] [doi:
       10.1136/bmj.f5980] [Medline: 24136633]
28.   Schmölzer GM, Roehr CC. Use of respiratory function monitors during simulated neonatal resuscitation. Klin Padiatr 2011
       Sep;223(5):261-266. [doi: 10.1055/s-0031-1275696] [Medline: 21630178]
29.   Garfield JM, Paskin S, Philip JH. An evaluation of the effectiveness of a computer simulation of anaesthetic uptake and
       distribution as a teaching tool. Med Educ 1989 Sep;23(5):457-462. [doi: 10.1111/j.1365-2923.1989.tb00902.x] [Medline:
       2677616]
30.   Schwid HA, Rooke GA, Ross BK, Sivarajan M. Use of a computerized advanced cardiac life support simulator improves
       retention of advanced cardiac life support guidelines better than a textbook review. Crit Care Med 1999 Apr;27(4):821-824.
       [doi: 10.1097/00003246-199904000-00045] [Medline: 10321676]
31.   Dweck C. Self-theories: Their Role in Motivation, Personality, and Development. Philadelphia, PA: Psychology Press/Taylor
       and Francis; 1999:1-195.
32.   Simulation-based training for health professionals. RETAIN Labs Medical Inc. Edmonton, Canada; 2020 Nov 28. URL:
       http://retainlabsmedical.com [accessed 2020-11-28]
33.   BioWare Corp. Edmonton, Canada; 2020 Nov 28. URL: https://www.bioware.com/games [accessed 2020-11-28]
34.   Ghoman SK, Schmölzer GM. The RETAIN Simulation-Based Serious Game-A Review of the Literature. Healthcare (Basel)
       2019 Dec 22;8(1) [FREE Full text] [doi: 10.3390/healthcare8010003] [Medline: 31877882]
35.   R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020
       Nov 28. URL: https://www.R-project.org [accessed 2020-11-28]
36.   Frank EHJ. Hmisc: Harrell Miscellaneous. R package version 4. 2020 Nov 28. URL: https://CRAN.R-project.org/
       package=Hmisc [accessed 2020-11-28]
37.   Revelle W. psych: Procedures for psychological, psychometric, and personality research. Evanston, IL: Northwestern
       University URL: https://CRAN.R-project.org/package=psych [accessed 2020-11-28]
38.   Mair P, Wilcox R. Robust statistical methods in R using the WRS2 package. Behav Res Methods 2020 Apr;52(2):464-488.
       [doi: 10.3758/s13428-019-01246-w] [Medline: 31152384]
39.   Pena E, Slate E. gvlma: Global validation of linear models assumptions. R package version 1. 2020 Nov 28. URL: https:/
       /CRAN.R-project.org/package=gvlma [accessed 2020-11-28]
40.   Johnson PO, Fay LC. The Johnson-Neyman technique, its theory and application. Psychometrika 1950 Dec;15(4):349-367.
       [doi: 10.1007/bf02288864]
41.   Johnson P, Neyman J. Tests of certain linear hypotheses and their applications to some educational problems. Statist Res
       Memoirs 1936;1:1-93.
42.   Long J. interactions: Comprehensive, user-friendly toolkit for probing interactions. R package version 1. URL: https://cran.
       r-project.org/package=interactions [accessed 2020-11-28]
43.   Maechler M, Rousseeuw P, Croux C. robustbase: Basic robust statistics R package version 0. URL: http://CRAN.
       R-project.org/package=robustbase [accessed 2020-11-28]
44.   Koller M, Stahel WA. Sharpening Wald-type inference in robust regression for small samples. Computational Statistics &
       Data Analysis 2011 Aug;55(8):2504-2515. [doi: 10.1016/j.csda.2011.02.014]
45.   Yohai V. High breakdown-point and high efficiency estimates for regression. The Annals of Statistics. URL: http://www.
       jstor.org/stable/2241331 [accessed 2020-11-28]
46.   Barré J, Michelet D, Truchot J, Cabon P, Tesniere A. Midwifery students' retention of learning after screen-based simulation
       training on neonatal resuscitation: a pilot study. BMJ STEL 2020 Apr 06. [doi: 10.1136/bmjstel-2019-000525]
47.   Padilha JM, Machado PP, Ribeiro A, Ramos J, Costa P. Clinical Virtual Simulation in Nursing Education: Randomized
       Controlled Trial. J Med Internet Res 2019 Mar 18;21(3):e11529 [FREE Full text] [doi: 10.2196/11529] [Medline: 30882355]
48.   Edwards D, Lampotang S. Computer-and web-based simulators and virtual environments. In Comprehensive Health care
       Simulation: Anesthesiology. 2020 (pp. 117-125). Springer, Cham 2020 Nov 28. [doi: 10.1007/978-3-030-26849-7_12]

49.   Cutumisu M, Lou N. The roles of generic and domain-specific mindsets in learning graphic design principles. Interactive Learning Environments 2020 (in press). [doi: 10.1080/10494820.2020.1857787]

Abbreviations

HCP: health care provider
NICU: neonatal intensive care unit
NRP: Neonatal Resuscitation Program
RETAIN: REsuscitation TrAINing
VIF: variance inflation factor

XSL•FO
RenderX

Original Paper

A Mobile Serious Game About the Pandemic (COVID-19 - Did You Know?): Design and Evaluation Study

Juliano De Souza Gaspar[1*], PhD; Eura Martins Lage[1*], PhD; Fernando José Da Silva[2*], PhD; Érico Mineiro[2*], PhD; Isaias José Ramos De Oliveira[1*], MSc; Igor Oliveira[3*]; Rayner Guilherme De Souza[1*]; Juan Rodrigues Oliveira Gusmão[1*]; Camila Fernanda Donadoni De Souza[3*]; Zilma Silveira Nogueira Reis[1*], PhD

[1]Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

[2]Departamento de Tecnologia de Design, Arquitetura e Urbanismo, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

[3]Curso de Design, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

[*]all authors contributed equally

Corresponding Author:
Juliano De Souza Gaspar, PhD
Faculdade de Medicina
Universidade Federal de Minas Gerais
Rua Alfredo Balena, 190
Belo Horizonte, 30130-100
Brazil
Phone: 55 313409 0000
Email: julianogaspar@gmail.com

Abstract

Background: No treatment for COVID-19 is yet available; therefore, providing access to information about SARS-CoV-2, the transmission route of the virus, and ways to prevent the spread of infection is a critical sanitary measure worldwide. Serious games have advantages in the dissemination of reliable information during the pandemic; they can provide qualified content while offering interactivity to the user, and they have great reach over the internet.

Objective: This study aimed to develop a serious game with the purpose of providing science-based information on the prevention of COVID-19 and personal care during the pandemic while assessing players' knowledge about COVID-19–related topics.

Methods: The study was conducted with the interdisciplinary collaboration of specialists in health sciences, computing, and design at the Federal University of Minas Gerais, Brazil. The health recommendations were grouped into six thematic blocks, presented in a quiz format. The software languages were based on the progressive web app development methodology with the Ionic framework, JavaScript, HTML5, cascading style sheets, and TypeScript (Angular). Open data reports of how users interact with the serious game were obtained using the Google Analytics application programming interface. The visual identity, logo, infographics, and icons were carefully developed by considering a selection of colors, typography, sounds, and images that are suitable for young audiences. Cards with cartoon characters were introduced at the end of each thematic topic to interact with the player, reinforcing their correct answers or alerting them to the need to learn more about the disease. The players' performance was assessed by the rate of incorrect and correct answers and analyzed by linear correlation coefficient over 7 weeks. The agile SCRUM development methodology enabled quick and daily interactions of developers through a webchat and sequential team meetings.

Results: The game "COVID-19–Did You Know?" was made available for free on a public university website on April 1, 2020. The game had been accessed 17,571 times as of September 2020. Dissemination actions such as reports on social media and television showed a temporal correspondence with the access number. The players' error rate in the topic "Mask" showed a negative trend ($r=–.83$; $P=.01$) over the weeks of follow-up. The other topics showed no significant trend over the weeks.

Conclusions: The gamification strategy for health education content on the theme of COVID-19 reached a young audience, which is considered to be a priority in the strategy of orientation toward social distancing. Specific educational reinforcement measures were proposed and implemented based on the players' performance. The improvement in the users' performance on the topic about the use of masks may reflect an increase in information about or adherence to mask use over time.

Introduction

Background

Two months after the first reports of COVID-19 in China, the first case was officially registered in Brazil [1,2]. At this time, the new coronavirus (SARS-CoV-2) had already been identified in more than 50 countries, accounting for 87,000 cases and 3000 confirmed deaths according to the World Health Organization (WHO) [3]. In November 2020, on the date of preparation of this paper, more than 52.4 million cases of COVID-19 and 1.2 million deaths due to the disease had already been registered worldwide [3]. In the same period, in Brazil, more than 5.8 million cases and 164,000 deaths by COVID-19 were registered [4].

In the absence of a short-term treatment or vaccine, providing access to information about SARS-CoV-2, the transmission route of the virus, and ways to prevent the spread of infection has been the focus of health strategies. To help countries prepare themselves to face the pandemic, the WHO has provided guidance and training on how to prevent and delay the transmission of the disease. Personal hygiene recommendations, such as washing hands frequently with soap and water, wearing a mask in public, avoiding handshakes, and social distancing whenever possible, continue to be widely disseminated in the population through various communication channels [5]. However, for such measures to be effective, in addition to government boosting, community awareness and engagement are considered to be critical factors to control disease [6].

The disease caused by SARS-CoV-2 is relatively mild in young adults, teenagers, and children [7]. Most people in these age groups, even when infected, are asymptomatic or oligosymptomatic; this creates concern regarding the potential of this population to transmit the disease, especially by direct contact with people in high-risk groups [8]. Also, physical and social distancing drastically reduced the opportunities for collective engagement among young people, causing psychological distress in many and leading to breaking of the distancing pacts that are so relevant at this time [9]. Another important consideration is the ease with which adolescents can use technology, which is vital to keep communication channels open and help adolescents inform and support each other in addition to sharing information with most of the community.

Digital health solutions can be a promising approach to address the spread of COVID-19; digital tools can effectively support institutions, facilitating the wide dissemination of information [10]. Communication during a pandemic must reach the target population in a timely fashion and provide clear, objective information. Combating misinformation and fake news about the origin, dissemination, and treatment of COVID-19 is a strategy that enables citizens to increase their adherence to the measures recommended during the crisis [11]. Thus, a good communication strategy avoids confusion and distrust, which can have negative consequences for individuals and society [11].

In the context of strict rules on social distancing, a serious digital game can offer significant advantages for the dissemination of information and learning because it does not require the user's physical presence, increases their interactivity with the content, and provides wide information coverage [12]. A serious game is characterized as a game in which the main purpose is not entertainment and fun [13,14]. Serious games can be powerful tools for the development and acquisition of new knowledge and skills by experienced users as well as by beginners [15].

Objective

The aim of this study was to develop and evaluate a serious game for mobile platforms to provide information about prevention of COVID-19 and personal care during the pandemic.

Methods

Study Design

The applied research in this study has an interdisciplinary profile between medicine, computer science, and design. In this study, the development, implementation, and evaluation of a serious game that addresses topics related to preventive measures and information about COVID-19 is presented. Teenagers are the target audience for this serious game; however, it can also be played by literate children and adults.

Theoretical Basis

A literature review on COVID-19, gamification, serious games, mobile apps, and e-learning was conducted. Concerning the topics on COVID-19 and prevention recommendations, it was decided to use only the information and recommendations available on the WHO website [3]. In addition to the initial research, periodic queries were made to the WHO website to obtain updates on recommendations and guidelines for action. A team of specialists, including physicians, professors, and medical students, reviewed and validated all content used in the serious game.

Learning Objectives

The learning content of this serious game was grouped into six topics that presented specific WHO recommendations for the population, with an emphasis on issues related to the daily lives of teenagers:

- Coronavirus: information about COVID-19 and vulnerable groups
- Mask: why and how to wear masks
- Take Care: transmission of COVID-19
- Cleaning: care for cleaning the home and tools for work and study
- Health: personal health, routines, and life habits
- Social: socializing with friends and school, how to shop, care outside the home

System Requirements

Based on the learning objectives and the identification of the target population, the system requirements were defined and a

search of specialized websites on the characteristics of frequently used devices was conducted. Items such as screen resolution and the amount of internet access via mobile phone, tablet, or desktop [16], as well as aspects such as reach, necessary network specifications, and data consumption were analyzed. The design team also considered specific characteristics for gamification and gameplay [15]. The most important requirements that guided the development are listed below.

- The game is accessible on different platforms (mobile and desktop web).
- It does not require registration to use.
- It does not use gaming platforms that require robust hardware, that is, it is possible to run it with simple processors and little memory.
- It consumes little internet browsing data.
- It maintains the player's score history,
- It enables sharing of the results of the phases on social networks.
- It enables users to view their personal and global ranking in the game.
- It offers information complementary to the questions of the game.
- It offers information about the project, the institution, and the team.

Design and Development

Agile SCRUM development methodology [17] was used in this project, and a multidisciplinary team was systematically gathered around the project. Weekly tasks were defined for each participant according to their qualifications. The tasks were developed during a cycle ("sprint") lasting 1 week [18]; rapid daily interactions were conducted through a webchat and weekly meetings with the entire team (on the web) to present the results for the week and define new tasks for the new sprint.

To create a multiplatform application, the progressive web app (PWA) software development methodology with the Ionic framework, using JavaScript, HTML5, cascading style sheets (CSS), and TypeScript (Angular) programming languages, was used. For the development, the Visual Studio Code tool [19] was used, and the code management was controlled with the Bitbucket tool [20] to enable the team to collaborate on codes and test them through the Git versioning system [21]. The Web Storage application programming interface (API) [22] was used to store information in the user's browser about the use and performance of the questions.

Open data reporting from the Google Analytics API was used to obtain information on how users interact with serious gaming [23]. Only information related to the use of a game is collected, such as the type of device, screen resolutions, clicks on right and wrong answers, how users accessed the game (via external links, social networks, search engine searches, news sites), and the time spent in the game. Even considering only very generic data, data retention by Google Analytics was configured for maximum storage of 26 months on its servers [24].

Communication and Diffusion of the Game

Complementary actions were implemented to diffuse the game. Among these, the creation of social network accounts on Facebook [25] and Instagram [26] are the most notable, in addition to sending emails to the population previously registered for the Faculty of Medicine of Universidade Federal de Minas Gerais (UFMG) newsletter [27]. Emails were also sent specifically to coordinators and teachers of elementary and local high schools.

Statistical Analysis

The data used in this analysis were retrieved from the Google Analytics weekly reports. Data such as the access number and the numbers of right and wrong answers were analyzed.

Continuous variables are represented as median values and categorical variables for absolute and relative values. Specifically, for the numbers of right and wrong answers, we calculated the hit rates and error rates for each topic aggregated by week. The Pearson correlation coefficient was used to assess the temporal variation over the weeks, and the Student t test was calculated to assess the significance of the results found by the coefficient. The significance used was $P<0.05$, and the calculations were performed using Google Sheets.

Line graphs were used to present the variation of error rates over the weeks evaluated, and the trend line was presented with the Pearson linear correlation coefficient (r) for topics with statistical significance. Sector charts and maps were also used to describe the categorical variables.

Ethical Aspects

This project does not use sensitive user data because the developed game does not involve user registration and does not have an associated database. The data used in the analysis of this study were generated by indirect reports (generated by Google Analytics) and with temporary data regarding access to the site. This project seeks to comply with all data usage laws in force in Brazil and Europe, including clarifying the indirect use of nonsensitive data to the user in an objective, clear, and timely manner. The project was developed and supported by the Center for Health Informatics of the Faculty of Medicine and the School of Architecture at UFMG.

Results

The game "COVID-19–Did You Know?" was made available for free on the web on April 1, 2020. Using the PWA methodology, the game was published on the server of the Faculty of Medicine of UFMG; submission to app stores was not required, and the game could be accessed directly on the website [28].

During the development process with the SCRUM methodology, 23 cycles lasting one week were performed. In each cycle, several independent tasks were assigned to the team members, including the literature review, definition of system requirements, quiz preparation, development of the game logic, interfaces, player ranking, final design, software testing, English and Spanish translations, and publication of the game on the web.

XSL•FO

RenderX

Learning Objectives

Learning objectives were defined based on the target population, and the information was grouped into the topics Coronavirus, Mask, Take Care, Cleaning, Health, and Social. Figure 1 shows the six topics with the number of questions created in each set.

Figure 1. Learning objective topics and number of questions for each topic.
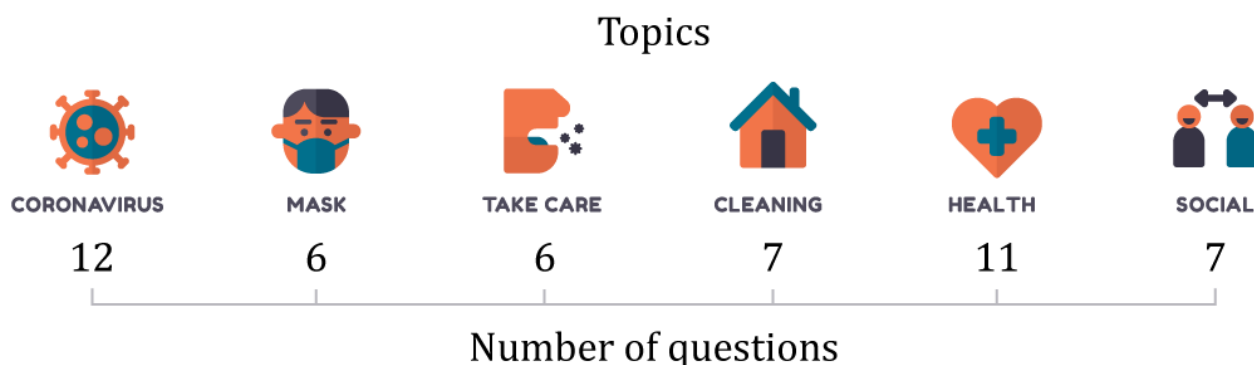
The Game

Design experts created the visual identity of the app. A dedicated logo and icons were developed (Figure 1); also, colors and typography were selected to meet the requirements defined in this project. The Attribution-Non-Commercial-Share-Equal 3.0 Brazil license [29] allows the free use of all images, infographics, and information pieces created for noncommercial purposes. Figure 2A shows the splash screen interface, which uses a minimalist concept with the game logo and the brands of the School of Architecture and the Faculty of Medicine of UFMG.

Figure 2. Interfaces of the initial screens of the game: Splash (A), Home (B), and Menu (C).

The main interface of the serious game, shown in Figure 2B, has buttons to access each set of questions on the defined topics. Along with each icon, a progress bar is displayed showing the player's evolution in the topic. A quick access bar was created at the top, providing the player with easy access to their ranking, changing the language of the game, and other information. In the upper left corner, a standard menu button gives access to all the resources and information in the game (Figure 2C). In the menu, it is possible to reset the score of the game, access the WHO website, learn more information about the topics, and learn more about the project and development team, in addition to consulting the complete Privacy Policy.

The game was developed in multiple languages to reach an audience beyond the borders of Portuguese-speaking countries. The second version of the app became available in September 2020; in addition to the Portuguese language, it added English and Spanish language options. In this version, all questions, answers, feedback, and infographics, as well as the Privacy Policy and complementary information, were translated into the new languages.

When the player selects a topic in the main interface (Figure 2B), they are presented with a set of questions; they can then answer or skip each question (Figure 3A). When the player answers the question, feedback about their right or wrong answer is presented to them (Figure 3B, 3C). Audiovisual effects, including icons, colors, and sounds, were used for the questions as well as for each type of feedback. All images have the alt attributes of the HTML5 language defined according to World Wide Web Consortium recommendations [30]; these attributes help to describe an image when it cannot be reproduced by the browser or even when the user employs reading aid software for the visually impaired.

Figure 3. Interface showing a question (A), feedback for the right answer (B), and feedback for the wrong answer (C).
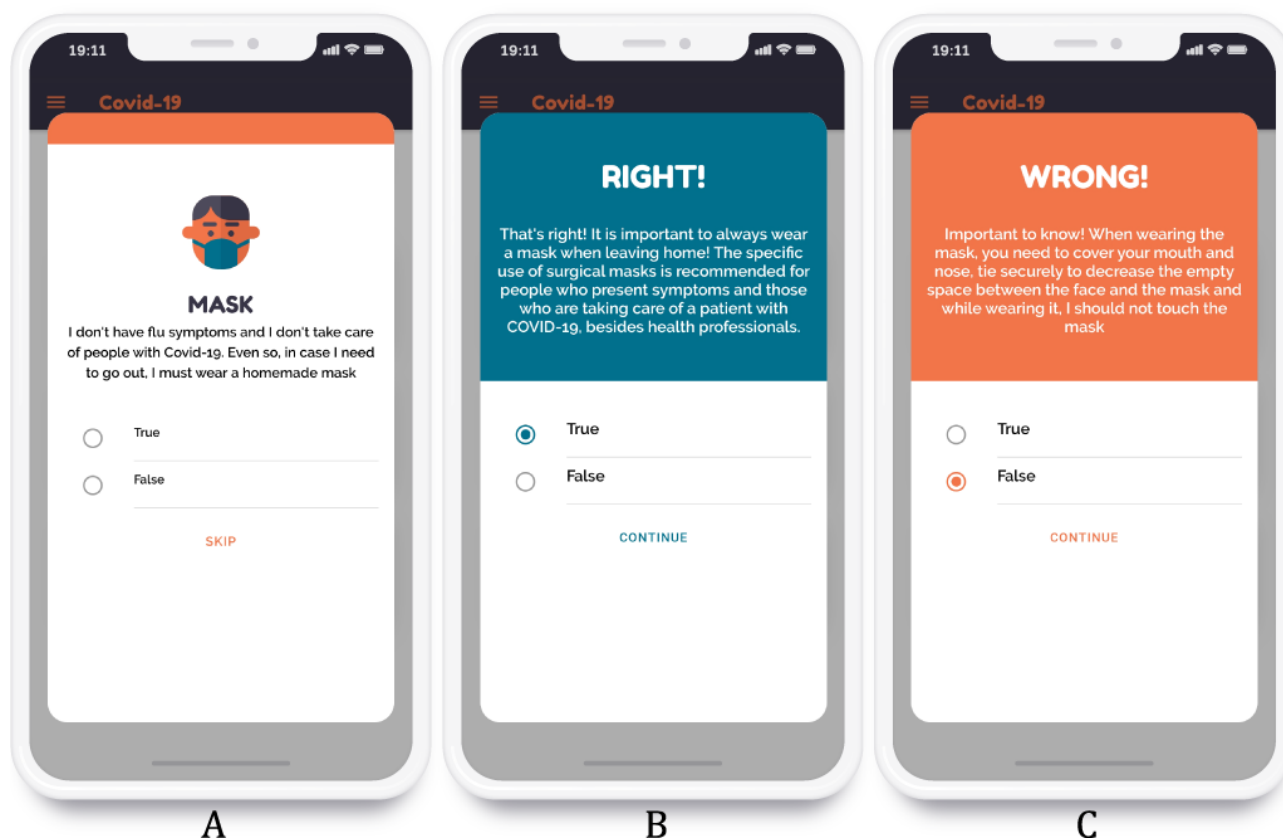
When the player finishes a topic, a card with a cartoon of a physician appears, congratulating the player for answering the questions and finishing that topic (Figure 4A) or asking them to pay more attention to the topic (Figure 4B). Female and male versions of the cartoon physician appear randomly. The user is also awarded different medals for each successfully completed topic (Figure 4C). The card also contains buttons the user can click to learn more about the topic, continue to the next topic, or post their performance on Facebook.

The Local Storage resource was employed to save user data locally, and Google Analytics was employed to collect nonsensitive user data. To comply with the General Data Protection Regulation (GDPR) in force in Europe and the corresponding Brazilian legislation, a notice about the use of the data is presented as soon as the app starts. This notice informs the user that no personal data are collected but that "cookies" and similar methods can be used to save their score [31]. The use of nonsensitive data is also requested, and the user is invited to read and accept the Privacy Policy containing detailed information (Figure 5-C). The complete text of the Privacy Policy [32] is provided in Multimedia Appendix 1.

Figure 4. Cards with results (A, B) and medal (C) for a topic.

Figure 5. Personal and global ranking (A), disclosure actions (B), and Privacy Policy (C) screens.

The player's score is stored in their browser using Web Storage technology; therefore, it is possible to show a progress bar on the main screen indicating the user's evolution by topic (Figure 2B). This also enables the user to compare their performance with that of other players in the global ranking interface (Figure 5A). Moreover, we implemented an algorithm so that when the user selects a topic they have already played, only questions they have not yet answered or for which they gave wrong answers are shown. The global ranking of right answers by topic was calculated from the reports generated by Google Analytics, which in this case collects the number of user clicks on correct or incorrect answers. Users can also opt to disseminate images from the game on social networks (Figure 5B).

The "Learn More" section presents infographics (Figure 6) with tips on how to proceed in situations related to each topic. Tips are provided for practicing physical activity, using masks, actions to take when leaving home, organizing work at home, hand washing care, and other personal care.

Figure 6. Infographics providing complementary information on the topics of the game.

Evaluation

This topic shows the data extracted from the reports generated by Google Analytics. Between April 1 and September 13, the game was accessed 17,571 times (Figure 7). In this period, the primary type of device used to access the game was smartphones (79.8%), followed by computers (19.2%) and tablets (1.0%). The average duration of the game use time in the analyzed period was 3 minutes and 34 seconds.

Figure 7. Distribution of game access between April and September 2020. Data are adapted from Google Analytics.
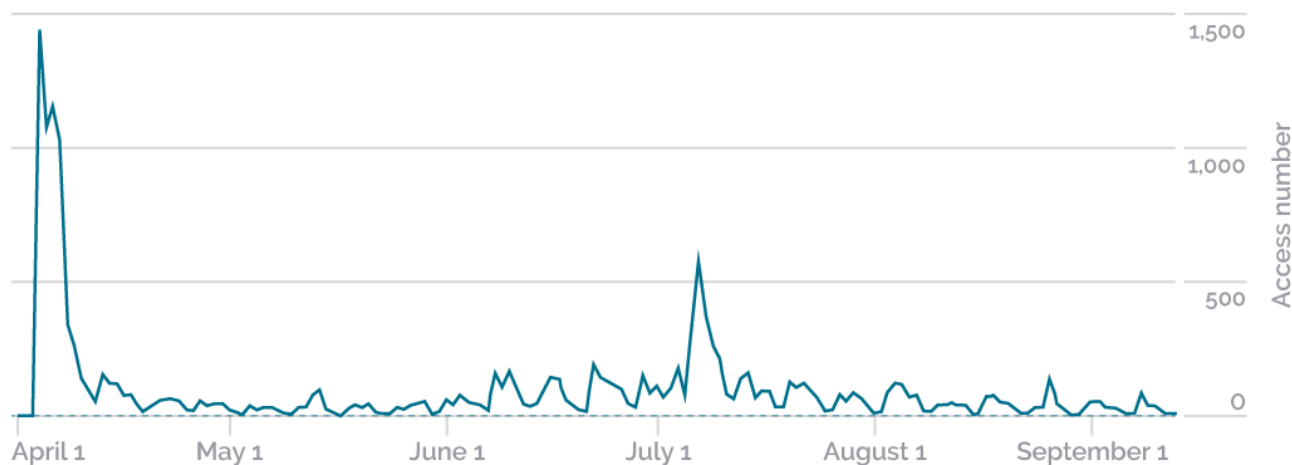
Figure 8 contains two graphs showing how the users access the game website and where they access it from. Most of the access types are direct access (79.3%), which includes situations in which the user types the address directly into the browser, accesses links saved in favorites or links in PDF files, and in some cases accesses links in emails (Figure 8A). Then, there is access through references (8.2%), which includes access via links on other sites, such as news reports and articles. Moreover, 7.5% of users arrived at the site by searching for keywords related to the game on searching sites, and 5.1% of users accessed it through links on social networks. Most users were from Brazil (98%). This period of analysis precedes the publication date of the multilanguage version (Figure 8B).

Figure 8. Distribution of game users by access channel (A) and by country (B) between April and September 2020. Adapted from Google Analytics.
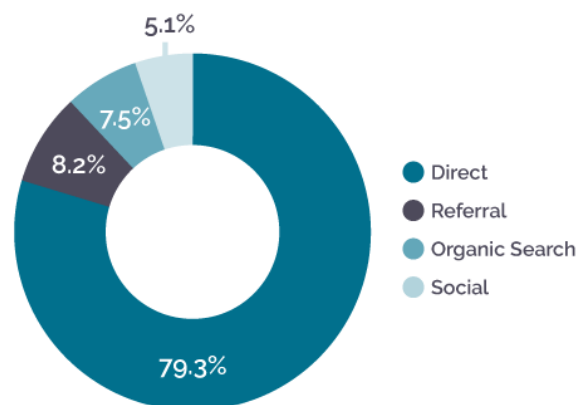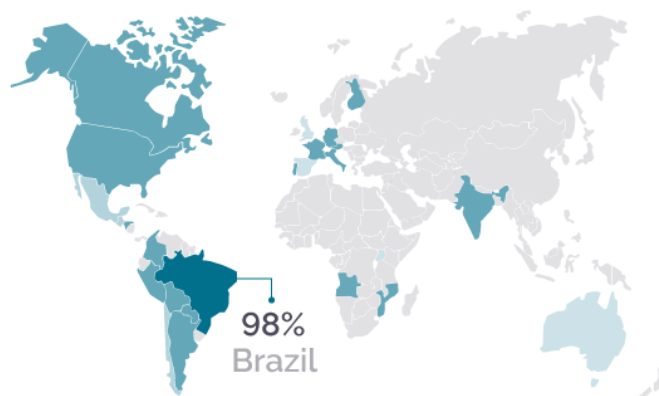
Adjustments to the software and the Google Analytics settings available from July 20, 2020, also enabled us to extract the users' correct and incorrect answer rates from the reports by question, both individually and grouped by topic. The correct answer rates for the topics varied between 69% and 89%, as shown in Figure 9.

Figure 9. Correct answer rate for each game topic between July 20 and September 13, 2020.

The error rates were grouped by topic and analyzed weekly (Table 1). A negative trend was observed only for the topic "Mask" ($r=-.83$), with a significant correlation ($P=.01$). The results for the other topics were not significant.

Table 1. Error rates by topic per week (July 20 to September 13, 2020).

Topic	Error rate per week (%)								r^a	P^b
	1	2	3	4	5	6	7	8		
Coronavirus	27.2	24.5	26.2	27.4	23.5	27.0	25.5	26.3	−.07	.87
Mask	17.0	17.3	19.2	18.8	14.8	13.7	12.2	12.2	−.83	.01
Take Care	33.0	29.8	30.5	32.4	30.0	31.5	28.8	33.7	−.002	.99
Cleaning	25.1	24.4	25.0	27.5	21.9	26.4	28.6	24.8	.25	.52
Health	18.1	17.4	18.6	18.0	17.9	17.9	19.2	19.1	.62	.08
Social	13.5	9.1	10.0	16.7	7.9	11.7	10.4	9.7	−.24	.53

[a]Pearson correlation coefficient.

[b]Student t test.

The graph in Figure 10 shows the error rates grouped by topic over the weeks evaluated. The linear trend ($r=-.83$) for the significant topic (Mask) indicated a downward trend in the error rates.

Figure 10. Error rates by topic grouped by week (July 20 to September 13, 2020). The linear trend for the Mask topic is significant ($r=-.83$).
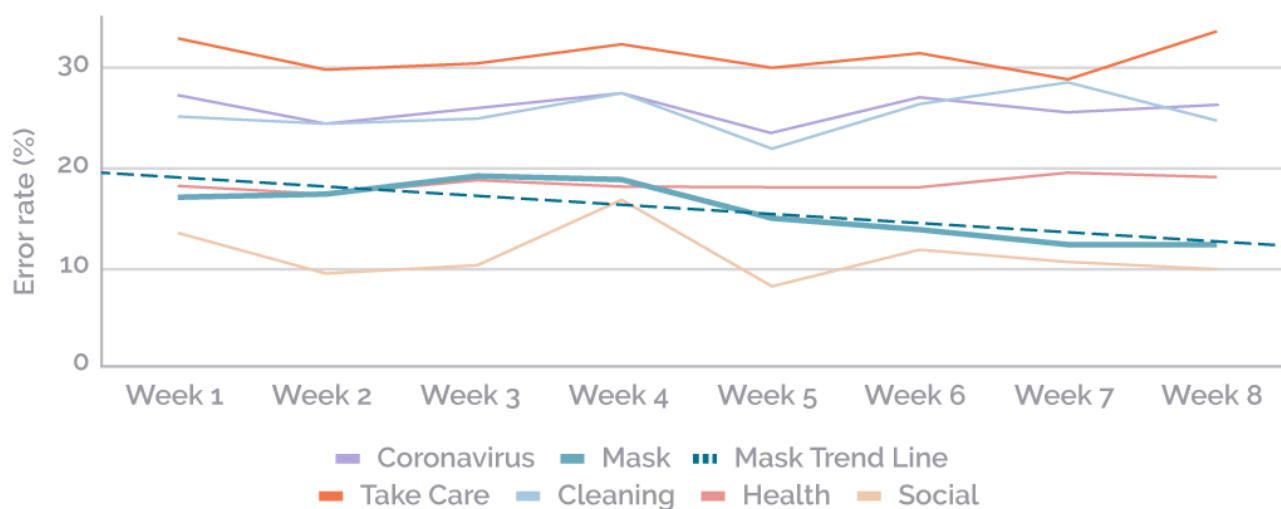
Publicizing Actions

Images and texts were produced by the design team related to the issues with the highest number of errors analyzed weekly (Figure 11). In this stage, images from the Unsplash [33] and Pexels [34] repositories that are offered free of charge by these platforms were searched for, selected, and used. From these images, graphic pieces were developed and published on the Facebook [25] and Instagram [26] accounts created for the project.

Figure 11. Examples of posts about the game on social networks.

Other actions to publicize the serious game included sending informative emails to the Faculty of Medicine of UFMG newsletter mailing list as well as specific emails to teachers and coordinators of elementary and high schools. The game was also offered as an educational strategy for homeschooling. We contacted 2476 schools by email between April and May and 1020 additional schools in September.

The lowest bounce rates in the Google Analytics reports were observed for users who accessed the game through newspaper links, interviews, and links from Google Classroom (38%).

Software Usability and Testing

The development team sought to meet the heuristic usability principles proposed by Nielsen and Molich [35]. We considered aspects such as offering simple dialogues, speaking the user's language, minimizing the user's memory overhead, maintaining consistent patterns of behavior and icons, offering continuous feedback to the user, providing demarcated exits with options to leave and return, providing shortcuts, avoiding error situations, offering clear error messages when necessary, providing an easy and intuitive interface, and offering help and clear documentation [36].

The serious game was subjected to empirical software development testing. White-box tests were performed to test game structures in specific parts of the development code for each component. Additionally, black-box tests were performed to validate the initially defined system requirements. Functional and nonfunctional items, such as performance, disclosure, acceptance, and release versions (alpha and beta), were evaluated. The observed inconsistencies were included in adjustments and incremental corrections in subsequent phases (sprints) of the development cycle.

Additionally, at least one visually impaired person played the game. His report was positive; he stated that he managed to use the game clearly with the aid of a specific reader for the visually impaired. He also reported the lack of information about the images related to the end of each topic. His feedback led to subsequent adjustments to improve the experience for visually impaired users, such as the inclusion of alternative text for these images.

Discussion

Principal Findings

The members of the multidisciplinary team, composed of physicians, designers, programmers, teachers, and students, were able to propose and develop solutions based on multiple aspects to implement the identified requirements, as recommended by Caserman and collaborators [15]. The weekly meetings were used as benchmarks to present the results of the week, evaluate the already implemented items, and propose new requirements for the development of an effective and attractive serious game.

The SCRUM development methodology proved to be efficient, enabling the division of the project into small increments; this methodology allowed for tests, rapid changes when necessary, and weekly deliveries to the end user. Functional software deliveries in a shorter period (1 week) generated greater customer satisfaction and provided a development environment with motivated individuals, as reported by Tobias and Spanier [37].

On January 30, 2020, the WHO declared that the outbreak of a disease caused by a new coronavirus, called COVID-19, constituted a Public Health Emergency of International Concern. This statement aimed to improve coordination, cooperation, and global solidarity to attempt to stop the spread of the new coronavirus. On March 11, 2020, the WHO characterized COVID-19 as a pandemic. Since the declaration of the outbreak and the characterization of the pandemic, the WHO has sought to inform the population about the health risks presented by COVID-19, considering that reliable information is as important as other protective measures. Well-informed people can make informed decisions and adopt positive behaviors to protect themselves and their families [38].

Much information about COVID-19 has been presented to the community; however, few initiatives are aimed at younger people. Thus, this audience was chosen because it has been more resistant to the recommendations of health authorities. Also, when teenagers are infected with SARS-CoV-2, they have few symptoms but can transmit the virus. We sought, through digital technology, to contribute to the rapid diffusion of information about SARS-CoV-2 and COVID-19, promoting changes in the attitude of the population [39].

Limited or insufficient health literacy has been associated with lower adoption of protective behaviors, such as vaccination, hand hygiene, and other self-care measures [40]. We avoided addressing specific issues in any country or region, avoiding major differences between the information provided by the WHO and the COVID-19 coping guidelines in each nation.

The use of PWA technology enabled the development of a hybrid app that combines the resources offered by browsers

with the advantages of using smartphones, as in traditional apps [41]. Through this technology, the app is compatible with most browsers; thus, it can be used on different devices, such as computers, tablets, and smartphones [42].

Using the Web Storage technology offered by HTML5 enabled local storage of player data, such as their position in the game, correctly answered questions, and personal ranking, more securely than using cookies [22]. Development with Web Storage ensures that data will never be transferred to the server by the browser, as can happen with cookies, ensuring greater adherence to user data security policies.

Regarding the visual identity of the app, different factors were considered during development. The seriousness of the topic addressed required the promotion of accessibility, usability, and information. Thus, we considered the use of typography that offers good legibility, adequate contrast between colors, and icons and images that are easy to recognize [43,44]. Moreover, to guarantee access to simpler devices and reduce noise in the information, the visual identity was developed with the aim to achieve a lean and responsive design [45]. The other factor to consider is the fun aspect of games, in which more subjective issues such as attractiveness, entertainment, and aesthetics are addressed, as proposed by Caserman and collaborators [15]. To achieve this target, we used vibrant colors, flashy and stylized typography, illustrations, sounds, and graphic unity between the elements.

During the 6 months of game analysis, during which 17,500 users accessed the game, there were 2 months (April and July) in which television and newspaper reports about the game were added to the actions of sending informative emails with the college newsletter. The game was accessed mostly by smartphones (79.8%), which indicates that it is a good choice to prioritize resources and functionalities for this type of access over access by computers and tablets.

The weekly monitoring of error answer rates enabled us to improve the text of some questions and answers. Similarly, the weekly assessment of answers grouped by topic enabled the team to develop complementary actions with images and informative text for publication on social networks targeting the topics with the most mistakes that week.

The finding of statistical significance in reducing the trend of the error rate for the topic "Mask" raised some questions among the team. Because this study is not a controlled clinical trial, it is not possible to say that the population is more informed about the importance and use of masks; however, this issue stands out and suggests the need for further studies.

The lowest bounce rates in the Google Analytics reports were obtained for access from newspaper links, interviews, and Google Classroom. Considering that the bounce rate indicates when a user opens and then closes the website (without interacting with it), the lower bounce rates in these segments may indicate that focusing on news channels and teachers is the most assertive way to acquire new users of this serious game.

Limitations

Because this is not a controlled study, the analysis permits limited inference that the results obtained are for a population of teenagers, which are the target population of this project. This is reinforced by the fact that the game does not have a user registry, is available on the internet, and is open to the community; therefore, the data analyzed in the current study do not enable inferences about who used the app.

Another limitation is that the sample used to assess the time series was seven weeks. Although statistical significance was found for one topic, evaluation for a longer period will be needed to state more safely that there is a decreasing error rate for that topic.

Conclusions

This study managed to comply with the proposed objectives of developing a serious game and making it available to young people, providing reliable information on topics related to the prevention of COVID-19. Also, the multidisciplinary profile of the team was able to bring reflections of its paradigms to the project; therefore, the game achieves compliance with the technical, legal, functional, and attractiveness requirements expected for a serious game. Extrapolating the initial requirements, we performed promotion and dissemination actions and increased the accessibility of the game, making it multilingual and accessible to people with visual impairments. This publication can provide an example not only to other students and teachers but also for those with future interest in the application of good practices in the development of serious game apps. We hope that this game will continue to combat misinformation on the topic of COVID-19 and expand the population's engagement in preventive measures against the disease.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Privacy Policy.

[PDF File (Adobe PDF File), 38 KB - games_v8i4e25226_app1.pdf]

References

1. Coronavirus disease (COVID-19) pandemic. World Health Organization. 2020. URL: https://www.who.int/emergencies/diseases/ [accessed 2020-08-23]
2. Primeiro caso de Covid-19 no Brasil permanece sendo o de 26 de fevereiro. Ministério da Saúde do Brasil. 2020 Feb 26. URL: https://www.gov.br/saude/pt-br/assuntos/noticias/primeiro-caso-de-covid-19-no-brasil-permanece-sendo-o-de-26-de-fevereiro [accessed 2020-08-23]
3. Coronavirus Disease (COVID-19) Dashboard. World Health Organization. 2020 Nov 3. URL: https://covid19.who.int/ [accessed 2020-11-13]
4. Painel Coronavírus. Ministério da Saúde do Brasil. 2020. URL: https://covid.saude.gov.br/ [accessed 2020-11-13]
5. World Health Organization. URL: https://www.who.int [accessed 2020-08-23]
6. Responding to community spread of COVID-19: interim guidance, 7 March 2020. World Health Organization. 2020 Mar 07. URL: https://apps.who.int/iris/handle/10665/331421 [accessed 2020-08-23]
7. Balasubramanian S, Rao NM, Goenka A, Roderick M, Ramanan AV. Coronavirus Disease 2019 (COVID-19) in Children - What We Know So Far and What We Do Not. Indian Pediatr 2020 May 15;57(5):435-442 [FREE Full text] [Medline: 32273490]
8. Huang L, Zhang X, Zhang X, Wei Z, Zhang L, Xu J, et al. Rapid asymptomatic transmission of COVID-19 during the incubation period demonstrating strong infectivity in a cluster of youngsters aged 16-23 years outside Wuhan and characteristics of young patients with COVID-19: A prospective contact-tracing study. J Infect 2020 Jun;80(6):e1-e13 [FREE Full text] [doi: 10.1016/j.jinf.2020.03.006] [Medline: 32283156]
9. Orben A, Tomova L, Blakemore S. The effects of social deprivation on adolescent development and mental health. Lancet Child Adolesc Health 2020 Aug;4(8):634-640 [FREE Full text] [doi: 10.1016/S2352-4642(20)30186-3] [Medline: 32540024]
10. Wetsman N. Effective communication is critical during emergencies like the COVID-19 outbreak. The Verge. 2020 Mar 04. URL: https://www.theverge.com/2020/3/4/21164563/coronavirus-risk-communication-cdc-trump-trust [accessed 2020-08-23]
11. Fagherazzi G, Goetzinger C, Rashid MA, Aguayo GA, Huiart L. Digital Health Strategies to Fight COVID-19 Worldwide: Challenges, Recommendations, and a Call for Papers. J Med Internet Res 2020 Jun 16;22(6):e19284 [FREE Full text] [doi: 10.2196/19284] [Medline: 32501804]
12. Suppan M, Gartner B, Golay E, Stuby L, White M, Cottet P, et al. Teaching Adequate Prehospital Use of Personal Protective Equipment During the COVID-19 Pandemic: Development of a Gamified e-Learning Module. JMIR Serious Games 2020 Jun 12;8(2):e20173 [FREE Full text] [doi: 10.2196/20173] [Medline: 32516115]
13. Djaouti D, Alvarez J, Jessel J, Rampnoux O. Origins of Serious Games. In: Ma M, Oikonomou A, Jain L, editors. Serious Games and Edutainment Applications. London, UK: Springer; 2011:25-43.
14. Rangel J. Jogo sério para ensino e prática de detecção de outliers. Thesis in Portuguese. Universidade Tecnológica Federal do Paraná. 2018. URL: http://repositorio.utfpr.edu.br:8080/jspui/handle/1/3933 [accessed 2020-12-15]
15. Caserman P, Hoffmann K, Müller P, Schaub M, Straßburg K, Wiemeyer J, et al. Quality Criteria for Serious Games: Serious Part, Game Part, and Balance. JMIR Serious Games 2020 Jul 24;8(3):e19037 [FREE Full text] [doi: 10.2196/19037] [Medline: 32706669]
16. Desktop vs Mobile vs Tablet Market Share Worldwide. StatCounter Global Stats. 2020. URL: https://gs.statcounter.com/platform-market-share/desktop-mobile-tablet [accessed 2020-08-23]
17. Sutherland J, Sutherland JJ. SCRUM. A arte de fazer o dobro do trabalho na metade do tempo. Rio de Janeiro, Brazil: Editora Sextante; Jan 01, 2019.
18. Knapp J. Sprint: How To Solve Big Problems and Test New Ideas in Just Five Days. New York, NY: Simon & Schuster; Mar 08, 2016.
19. Code Editing. Redefined. Visual Studio Code. 2020. URL: https://code.visualstudio.com [accessed 2020-08-23]
20. Bitbucket - The Git solution for professional teams. Atlassian. 2020. URL: https://bitbucket.org/product [accessed 2020-08-23]
21. Conheça o Git: tutoriais, fluxos de trabalho e comandos Git. Atlassian. 2020. URL: https://www.atlassian.com/br/git [accessed 2020-08-23]
22. HTML Web Storage API. W3Schools. 2020. URL: https://www.w3schools.com/html/html5_webstorage.asp [accessed 2020-08-23]
23. Gabriel JR, Passos P, Vanz SDS, Pavão C, Rocha RD, Borges E. Acesso aberto a dados de pesquisa no Brasil: Dataverse: monitoramento do DataVerse com o Google Analytics. CEDAP. 2019. URL: https://cedap.ufrgs.br/jspui/handle/20.500.11959/1203 [accessed 2020-08-23]
24. Google. Como ajudar os usuários a obedecer à Lei Geral de Proteção de Dados (LGPD). Ajuda do Google Ads. 2020. URL: https://support.google.com/google-ads/answer/9943919?hl=pt-BR [accessed 2020-08-23]
25. Covid 19 Você Sabia? Facebook. 2020. URL: https://www.facebook.com/covid19ufmg [accessed 2020-08-23]
26. Instagram. COVID-19 Você Sabia? Instagram. 2020. URL: https://www.instagram.com/covid19ufmg [accessed 2020-08-23]
27. Portal de Notícias da Faculdade de Medicina da UFMG. URL: https://www.medicina.ufmg.br/noticias [accessed 2020-08-23]

28.   Covid-19. Você sabia? A serious game. Faculdade de Medicina da UFMG. URL: https://site.medicina.ufmg.br/covid/home
       [accessed 2020-08-23]
29.   CCommons. Atribuição-NãoComercial-CompartilhaIgual 3.0 Brasil (CC BY-NC-SA 3.0 BR). Creative Commons. URL:
       https://creativecommons.org/licenses/by-nc-sa/3.0/br [accessed 2020-08-23]
30.   Web Content Accessibility Guidelines (WCAG) Overview. Web Accessibility Initiative (WAI). URL: https://www.w3.org/
       WAI/standards-guidelines/wcag [accessed 2020-08-23]
31.   Bardini J. Saiba como o GDPR afeta a política de coleta de dados do Google Analytics. Marketing Analítico. 2020. URL:
       https://marketinganalitico.com.br/
       saiba-como-o-gdpr-afeta-politica-de-coleta-de-dados-do-google-analytics-e-como-nao-perder-dados/ [accessed 2020-08-23]
32.   Política de Privacidade. Faculdade de Medicina da UFMG. URL: https://site.medicina.ufmg.br/covid/privacidade [accessed
       2020-12-15]
33.   Unsplash. URL: https://unsplash.com [accessed 2020-08-23]
34.   Pexels. URL: https://www.pexels.com/pt-br [accessed 2020-08-23]
35.   Nielsen J, Molich R. Heuristic evaluation of user interfaces. In: CHI '90: Proceedings of the SIGCHI Conference on Human
       Factors in Computing Systems. 1990 Mar Presented at: SIGCHI Conference on Human Factors in Computing Systems;
       1990; New York, NY p. 249-256. [doi: 10.1145/97243.97281]
36.   Davids MR, Chikte UME, Halperin ML. An efficient approach to improve the usability of e-learning resources: the role
       of heuristic evaluation. Adv Physiol Educ 2013 Sep;37(3):242-248 [FREE Full text] [doi: 10.1152/advan.00043.2013]
       [Medline: 24022770]
37.   Tobias G, Spanier AB. Developing a Mobile App (iGAM) to Promote Gingival Health by Professional Monitoring of
       Dental Selfies: User-Centered Design Approach. JMIR mHealth uHealth 2020 Aug 14;8(8):e19433 [FREE Full text] [doi:
       10.2196/19433] [Medline: 32795985]
38.   Folha informativa COVID-19 - Escritório da OPAS e da OMS no Brasil. Organização Pan-Americana da Saúde. 2020.
       URL: https://www.paho.org/pt/covid19#contagio [accessed 2020-08-23]
39.   Bai Y, Yao L, Wei T, Tian F, Jin D, Chen L, et al. Presumed Asymptomatic Carrier Transmission of COVID-19. JAMA
       2020 Apr 14;323(14):1406-1407 [FREE Full text] [doi: 10.1001/jama.2020.2565] [Medline: 32083643]
40.   Castro-Sánchez E, Chang PW, Vila-Candel R, Escobedo AA, Holmes AH. Health literacy and infectious diseases: why
       does it matter? Int J Infect Dis 2016 Feb;43:103-110 [FREE Full text] [doi: 10.1016/j.ijid.2015.12.019] [Medline: 26751238]
41.   Biørn-Hansen A, Majchrzak T, Grønli TM. Progressive Web Apps: The Possible Web-native Unifier for Mobile Development.
       In: Proceedings of the 13th International Conference on Web Information Systems and Technologies - Volume 1: WEBIST,.
       2017 Presented at: 13th International Conference on Web Information Systems and Technologies; April 25-27, 2017; Porto,
       Portugal p. 344-351. [doi: 10.5220/0006353703440351]
42.   Waranashiwar J, Ukey M. Ionic Framework with Angular for Hybrid App Development. Int J New Technol Res 2018
       May;4(5):01-02 [FREE Full text]
43.   Dul J, Weerdmeester B. Ergonomia prática 2nd ed. São Paulo, Brazil: Edgard Blucher; 2012.
44.   Preece J. Design de Interação: além da interação homem-computador. Quezon City, Philippines: Bookman; 2013.
45.   de Moraes A. Avisos Advertências e Projeto de Sinalização: Ergodesign informacional. Rio de Janeiro, Brazil: Rio Books;
       2002.

Abbreviations

API: application programming interface
CSS: cascading style sheets
GDPR: General Data Protection Regulation
HTML5: Hypertext Markup Language revision 5
PWA: progressive web app
UFMG: Universidade Federal de Minas Gerais
WHO: World Health Organization

XSL•FO
RenderX

Original Paper

Web-Based Virtual Learning Environment for Medicine Administration in Pediatrics and Neonatology: Content Evaluation

Alayne Larissa Martins Pereira[1], MSc; Casandra Genoveva Rosales Martins Ponce Leon[1], MSc; Laiane Medeiros Ribeiro[1], PhD; Guilherme Da Costa Brasil[2], MSc; Karen Karoline Gouveia Carneiro[1], CNS; Géssica Borges Vieira[3], MSc; Yuri Gustavo De Sousa Barbalho[1], CNS; Izabel Cristina Rodrigues Da Silva[1], PhD; Silvana Schwerz Funghetto[1], PhD

[1]Universidade de Brasília, Brasília Distrito Federal, Brazil

[2]Centro Universitário do Distrito Federal, Brasília Distrito Federal, Brazil

[3]Hospital Universitário Dr. Miguel Riet Corrêa Júnior, Rio Grande, Rio Grande do Sul, Brazil

Corresponding Author:
Alayne Larissa Martins Pereira, MSc
Universidade de Brasília
Centro Metropolitano, Ceilândia Sul
Brasília Distrito Federal
Brazil
Phone: 55 061 996118662
Email: alayne_larissa@hotmail.com

Abstract

Background: Worldwide, patient safety has been a widely discussed topic and has currently become one of the greatest challenges for health institutions. This concern is heightened when referring to children.

Objective: The goal of this study was to develop a virtual learning environment for medication administration, as a tool to facilitate the training process of undergraduate nursing students.

Methods: Descriptive research and methodological development with a quantitative and qualitative approach were used with stages of design-based research as methodological strategies. For the development of the virtual environment, 5 themes were selected: rights of medication administration, medication administration steps, medication administration routes, medication calculation, and nonpharmacological actions for pain relief. After development, 2 groups—expert judges in the field of pediatrics and neonatology for environment validation and undergraduate nursing students for the assessment—were used to assess the virtual learning environment. For the validation of the virtual learning environment by expert judges, the content validity index was used, and for the evaluation of the students, the percentage of agreement was calculated.

Results: The study included 13 experts who positively validated the virtual environment with a content validity index of 0.97, and 26 students who considered the content suitable for nursing students, although some adjustments are necessary.

Conclusions: The results show the benefit of the virtual learning environment to the training of nursing students and professional nurses who work in health care. It is an effective educational tool for teaching medication administration in pediatrics and neonatology and converges with the conjectures of active methodologies.

Introduction

Worldwide, patient safety has been a widely discussed topic, and has currently become one of the greatest challenges for health institutions. It is understood as the decrease, to an acceptable minimum, of the risk of unnecessary harm associated with health care, and although there are recommendations and progress in international literature, scientific knowledge must become incorporated into care practice [1].

The National Patient Safety Program, aligned with the objectives of the World Alliance for Patient Safety of the World Health Organization launched 6 patient safety protocols focusing on

the most problematic areas of patient safety, among them improving safety in the prescription, use, and administration of medicines [2].

The concern with patient safety is heightened when referring to children, since they are a vulnerable part of society [3]. The quality of care related to patient safety in nursing is connected to the entire care process, thus, the use of intelligent technologies and the use of standardization can provide greater safety for patients [4].

Communication between professionals, parents, and patients is a factor in pediatric hospitalization, and the consequences of communication failures can generate significant damage to patients and decrease the effectiveness and quality of the care provided. Thus, in pediatric patient care, the prevention of errors, the analysis of the factors that led to an error, aims to implement measures, and improvements contribute to patient safety [5].

It was assumed that the development of a virtual learning environment on medication administration in neonatology and pediatrics would contribute to the knowledge gain of nursing students, and it was postulated that the use of a virtual learning environment is an adequate educational strategy to promote

knowledge about pediatric medication administration, encouraging active and motivating learning.

Design-based research involves a new interventionist methodology that seeks to incorporate theoretical aspects of education research with educational practice and offers a promising approach to advance theoretical understanding; although theory alone will not solve all the problems encountered in practice, it can help explain how and why problems can be solved and can help students get better results [6].

Thus, based on the principles of design-based research, to contribute to continued professional education and student learning and to complement the principles of sustainability and patient safety, we developed, validated, and evaluated a virtual learning environment.

Methods

We used descriptive research, methodological development with a quantitative and qualitative approach, and the stages of design-based research as methodological strategies [7]. The phases of the research are described in Figure 1.

Figure 1. Research phases. VLE: virtual learning environment.
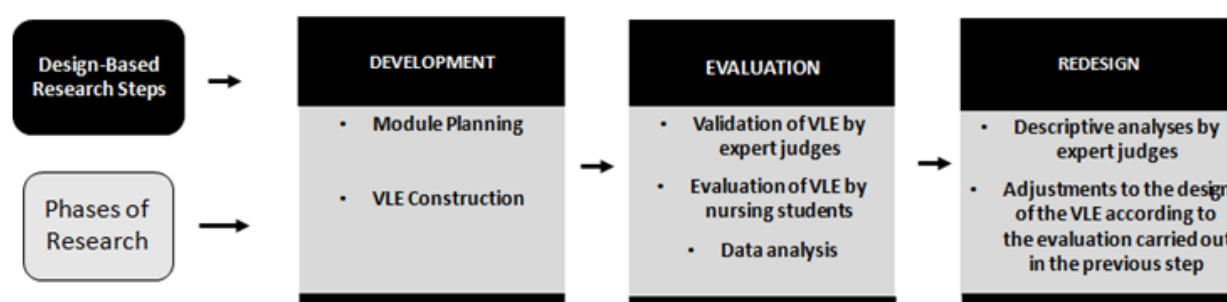
For the selection of the themes used in the virtual learning environment, main themes suggested by professionals in a previously conducted study [8] were taken into account, and 5 main themes were then selected: rights of medication administration, medication administration steps, medication administration routes, medication calculation, and nonpharmacological actions for pain relief.

After the construction of the virtual learning environment, 2 groups of participants, with different profiles and purposes, were included: experts in the field of pediatrics and neonatology (teachers and researchers) and final users (undergraduate students of nursing). After construction of the virtual environment, the content was validated by the experts and then evaluated by the students.

Students regularly enrolled in nursing at a public university were considered (convenience sample) and 112 students were invited; 26 students completed the evaluation by answering all the requested instruments corresponding to the final sample.

Data collection with students was carried out from June 2019 to August 2019. Participants were contacted by email and in person. Nursing students, through a questionnaire on Google

Forms, filled out a 24-item Likert scale developed in a previous study [9] and adapted for this research. The students judged the virtual environment in terms of interaction and stimulus (6 items); interest and motivation to learn (3 items); dedication, discipline and time management (3 items); communication tools (3 items); didactic material (7 items); and the student's role in the learning process (2 items). Space was reserved for suggestions and additional comments. Quantitative data were calculated using mean and standard deviation using SPSS (version 20; IBM Corp). The percentage of agreement was also calculated, individually and by categories—80% agreement among users was considered as a positive evaluation.

The validation of the virtual learning environment was executed by a committee of specialists with evident knowledge in pediatrics or knowledge of active technologies and with the ability to evaluate the content of the virtual environment. A total of 59 specialists were invited, and 13 completed all the steps and answered all the requested instruments, corresponding to the final sample. To define the specialists, the Fehring Validation Model [10] was used, which consists of criteria used to classify the expertise of professionals in matters covered in

educational technology. Nurses with specialized knowledge in pediatrics, neonatology, or active methodologies were included.

The experts answered a structured questionnaire focusing on the nature of the objectives, content, relevance, and environment. In the last section of the questionnaire, a blank space was reserved for the experts to add suggestions and comments (they were able to make qualitative collaborations and contribute to the improvement of the virtual learning environment). The content validity index was used, which measures and evaluates whether the proportion or percentage of experts are in agreement on certain aspects of the instrument and its items (each item individually and then the instrument as a whole). The questionnaire used a Likert-type scale with a score of 1 to 4 [11]. The content validity index of each item was calculated, and the content validity index of each category was calculated. Finally, the mean general content validity index of the virtual learning environment was calculated. For this study, a content validity index of 0.90 was considered as the experts judging the questions to be valid, based on Polit and Beck [12] who establish this value as standard to guarantee the excellence of a scale.

To measure the gain of knowledge, Worral [13] proposed the application of a pretest before the educational intervention, to identify the preacquired knowledge and a posttest to identify the gain of knowledge (retention capacity and memory) as an educational proposal. For the pretest, it was decided to carry out a questionnaire with 13 questions referring to the contents of the virtual environment, and those same questions were asked again in the posttest. The questionnaire contained true or false questions and underwent validation with 3 pediatric specialists, before being made available in the virtual environment. The students' responses to this questionnaire generated data on knowledge gain. With the Wilcoxon statistical test, we compared the outcome of the scores (pretest and posttest).

The research project was approved by the research Ethics Committee of the University of Brasília - Faculdade de Ceilândia, obtaining an opinion favorable to its realization under the protocol no. 3,203,397 and CAAE no. 08880519.3. 0000.8093.

The virtual learning environment for medication administration in pediatrics and neonatology was made available online with free access. The platform used to host the site was Heroku, a cloud computing service, and the reason for choosing this environment was the possibility of inserting images and videos created exclusively for a virtual learning environment.

The environment consisted of approximately 77 pages (HTML language), one of which was the main menu, and 8 were considered submenus of the themes of the modules. After registering, the user received an email, automatically generated by the platform as soon as the administrator approved participation. When accessing the virtual learning environment with login and password, the user was automatically redirected to the pretest page.

After answering the pretest questions, the user clicked on the *save answers* button and was directed to the main menu (Figure 2). This screen contained a button on the right side that accessed the 5 modules (Table 1), a button to check performance, an indication of the content already viewed, and a button for tutorial and credits. On the access screen for each module, it was possible to observe different interactions. Each circle generated a clickable link that opened a text box. Modules contained complementary material along with the reference used to write the content of that module. The links of the complementary material, when clicked, were directed to the publications.

At the end of the fifth module, the user was automatically redirected to the posttest page to answer the same 13 questions. After saving the answers, the user was directed to the home page and with the check performance button, could review which questions had been correct and which questions had been wrong. For the incorrect questions, a template showing correction and justification appeared.

Figure 2. Main menu of the virtual learning environment.
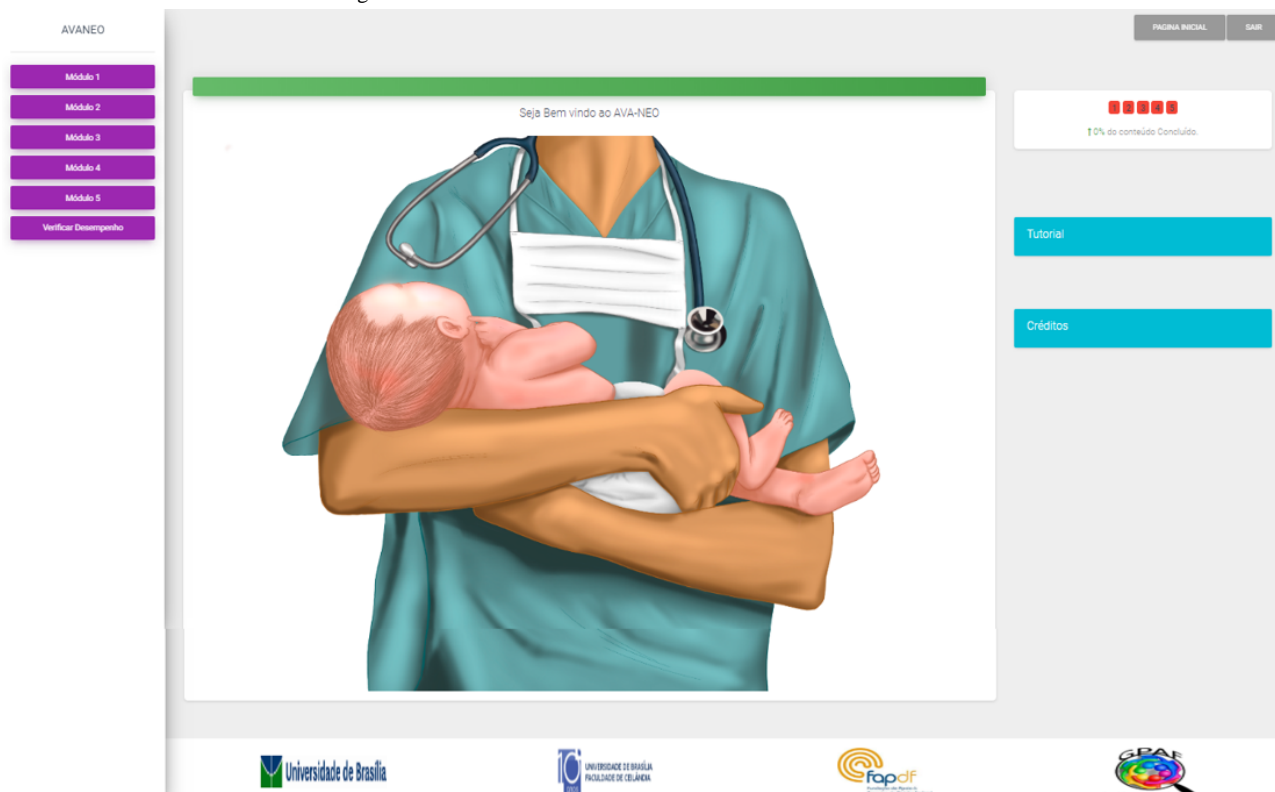
Table 1. Virtual learning environment modules.

Module	Features
Nine rights of medication administration	Right patient, right drug, right route, right time and frequency, right dose, right documentation, right education and information, right pharmaceutical form—each circle opens a text box; complementary materials
Medication administration steps	3 clickable boxes open content on absorption, distribution, and elimination, complementary materials
Medication administration routes	4 videos
Medication calculation	Formulas for calculating medication doses, complementary materials
Nonpharmacological actions for pain relief	Images with accompanying text, complementary materials

Results

Validation

The sample was composed of 13 nursing professors; 69% (9/13) had greater than 10 years of professional experience, and 46% (6/13) had a doctorate, and 46% (6/13) has a master's degree.

The experts' ages ranged from 25 to 55 years, and 31% (8/13) of the sample were aged 25 to 35 years; 85% (11/13) of the experts were female. Regarding having experience in the area of active teaching methodologies, 100% (13/13) answered "yes." Of the 13 experts, 54% (7/13) had publications in the area of pediatrics and neonatology.

Table 2 describes the assessments of the criteria analyzed and judged by the expert judges. No judge rated the analysis criteria as "Inadequate." The content validity index by category ranged from 0.91 to 1.0. The total content validity index reached was 0.97, which successfully validated the content.

Table 2. Responses from experts and content validity index results.

Items	Inadequate, n	A little adequate, n	Quite adequate, n	Very adequate, n	CVI[a]
Nature of the objectives					
Was the purpose of the VLE[b] clear? (described in the Tutorial tab of the platform)	0	0	3	10	1.0
Are the verbs chosen accurately in terms of expected behavior?	0	0	5	8	1.0
Are the objectives consistent with the content presented?	0	0	3	10	1.0
Content					
Does the content meet the theme and objectives proposed?	0	1	5	7	0.92
Is the content updated and does it contain correct information (sources and references)?	0	2	1	10	0.85
Are the texts easy to read?	0	1	2	10	0.92
Is the writing style compatible with the level of knowledge of the academics?	0	1	1	11	0.92
Is the number of modules sufficient (adequate division)?	0	1	6	6	0.92
Is the sequence of the modules adequate (clear and well-structured information)?	0	1	4	8	0.92
Relevance					
Do the items cover important aspects of nursing practice in the Administration of Medicines in Neonatology?	0	0	2	11	1.0
Does (the) VLE contribute to learning and knowledge acquisition?	0	0	3	10	1.0
Ambiance					
Is the VLE suitable for presenting the content?	0	0	1	12	1.0
Does VLE encourage student autonomy?	0	1	3	9	0.92
Do VLE resources offer learning situations?	0	0	4	9	1.0

[a]CVI: content validity index.

[b]VLE: virtual learning environment.

Evaluation

Of the 26 students who participated in the evaluation, 81% (21/26) were aged 19 to 24 years, 62% (16/26) are female and 38% (10/26) are male. The majority of students (17/26, 65%) were in the seventh semester (of 10) in undergraduate nursing.

Of the 24 items that were analyzed, 22 items and 5 of the 6 categories reached the stipulated value of 80% concerning percentage of agreement (Table 3). The categories *interaction and stimulus*, *instructional material*, and *the student's role in*
the learning process had the best percentage of agreement values (96% agreement). The category *communication tools* obtained 78% percent agreement, which was close to the stipulated value, indicating the need for improvements regarding the exchange of information between users and the administrator of the virtual environment. The mean Cronbach α was 0.88, indicating a high degree of reliability and consistency; users considered the content of the virtual learning environment suitable for nursing students, although some adjustments are necessary, for example, communication tools, which obtained the lowest percentage of agreement.

Table 3. Criteria evaluated by the users (n=26).

Item	Rating, mean (SD)	Agreement, %
Interaction and stimulation	**4.70 (0.47)**	**96**
The VLE[a] is easy to interact with and arouses your interest in the subject	4.73 (0.45)	100
The environment enables learning situations	4.69 (0.62)	92
VLE allows you to freely browse through content	4.69 (0.68)	96
The activities are relevant and meet the objectives proposed	4.77 (0.43)	100
Access to modules is easy	4.73 (0.53)	96
The VLE instigates/invites change in behavior and attitude	4.62 (0.64)	92
Interest and motivation to learn	**4 .59 (0.62)**	**93**
The VLE NEO-PED[b] motivates the students to learn about medication administration in neonatology and pediatrics	4.81 (0.40)	100
The VLE presents complementary readings and notes on the course content	4.50 (0.76)	92
The VLE encourages the desire to study beyond the content presented	4.46 (0.71)	88
Dedication, discipline, and time management	**4 .55 (0.66)**	**92**
During the access to the VLE, time was organized for online course activities	4.38 (0.80)	88
During access to the VLE, there was self-discipline for online teaching	4.62 (0.64)	92
During the access to the VLE, there was access to the modules with the regularity proposed	4.65 (0.56)	96
Communication tools	**4 .22 (1.04)**	**78**
The environment encourages the exchange of information with colleagues	4.19 (1.10)	77
The links provided are relevant to the learning content	4.62 (0.64)	92
An e-mail was used to answer questions about the content	3.85 (1.38)	65
Courseware	**4 .71 (0.56)**	**96**
The VLE is explanatory and easy to understand	4.69 (0.55)	92
The content is well divided between the modules	4.69 (0.55)	96
Information is presented in a logical and coherent way	4.65 (0.56)	96
The writing style is easy to understand	4.69 (0.55)	96
The colors and size of the letters are adequate	4.73 (0.67)	96
The images contribute to learning	4.85 (0.37)	100
The media are correlated with the content and provide a complement to the texts	4.73 (0.67)	96
Role of the student in the learning process	**4 .63 (0.55)**	**96**
You consider having freedom in building your knowledge	4.54 (0.58)	96
You believe you are responsible for your learning process	4.73 (0.53)	96

[a]VLE: virtual learning environment.

[b]NEO-PED: neonatology and pediatrics.

The results of the comparison between pretest and posttest correct answers are shown in the table below (Table 4). The pre–post difference in number of correct answers was statistically significant (P=.02).

Figure 3 likewise demonstrates the significant increase, corroborating the purpose of this study, and that students obtained a considerable gain of knowledge.

Table 4. Comparison of knowledge gain before and after use of the virtual environment.

Knowledge gain	Score, median (IQR)	P value
Pretest	75 (66-75)	.02
Posttest	75 (75-83)	

XSL·FO

RenderX

Figure 3. Knowledge gain before and after use.

Discussion

General
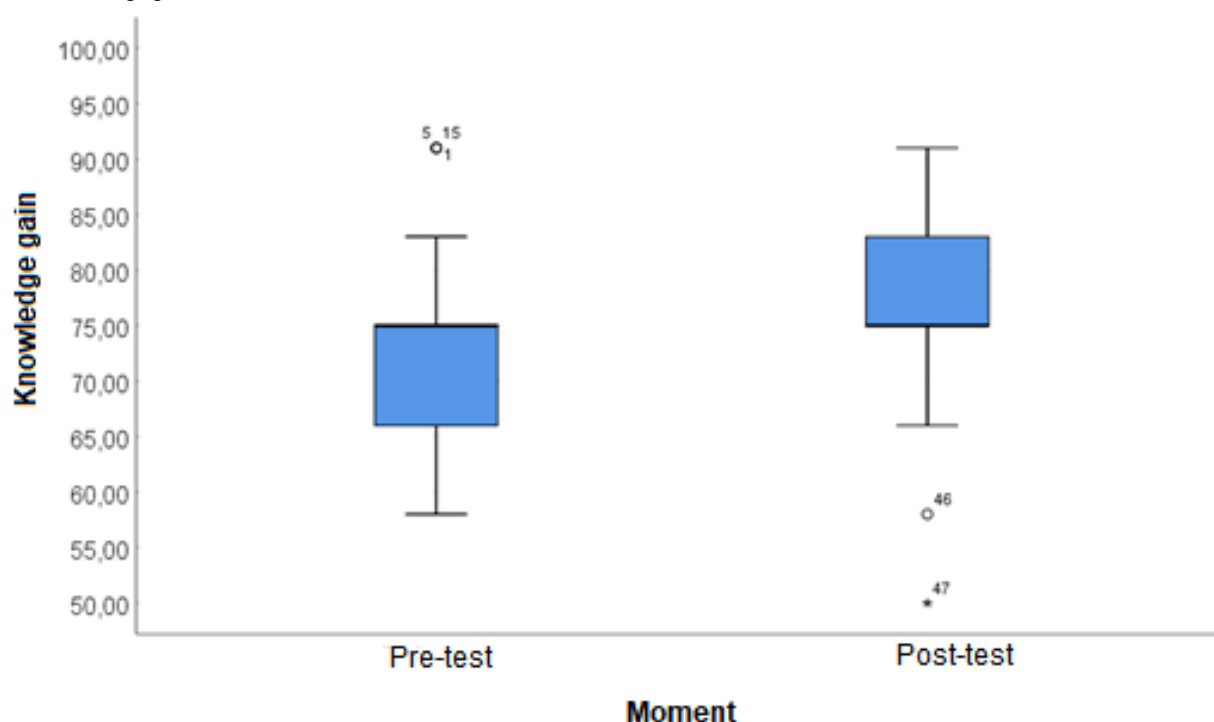
For the last stage, according to design-based research, there is a refinement of the intervention or re-design, in which adjustments are made to the design of the virtual learning environment according to the assessment that was made in the previous stage. At the end of the instrument, the expert judges had the opportunity to provide suggestions, changes, or opinions on the virtual environment. These opinions served as the basis for the changes in the virtual learning environment as a whole and not just in the content. Most of the comments emphasized the importance of virtual learning environment and its quality and excellence in this production. The suggestions were accepted, and changes were made as proposed by the experts.

Discussion

General

Through a bibliographic survey on the topic of medication administration in pediatrics and neonatology, it was possible to perceive the scarcity of research related to the use of virtual learning environments focused on this theme. Our objective was to develop a virtual learning environment with the theme of medication administration, as a facilitating tool for the training of undergraduate nursing students.

To refine the intervention of this study, descriptive analyses of the expert judges' opinions were used, as suggested at the end of the questionnaire. It should be noted that the judges' suggestions were analyzed, and the contents were modified, taking into consideration the target audience of this study.

According to the literature, electronic materials that provide information that contributes to the construction of knowledge are resources that support learning, especially for young students [14].

Virtual learning environments have been placed as innovative strategies that provide paradigm shifts for health care professionals; their development and use provide new educational possibilities to be explored both in universities and in continuing education [15]. Developing care techniques in safe environments for patients and professionals has become a priority for health systems, mainly in the development of strategies to prevent failures that affect good hospital practices [16]. By providing digital education, virtual environments have the potential to promote health education to the nursing team and are thus being used as an alternative or complement to traditional education for health professionals [17].

Measuring knowledge gained with the virtual environment showed an increase in correct answers in the posttest in comparison to the number of correct answers in the pretest. This finding proves that the online learning strategy can be an educational teaching resource to be used in the training of nurses or as a permanent education resource.

The use of educational technologies makes important contributions to the nursing teaching process, as they favor the enrichment of the academic's cognitive structure by enabling continuous access to information, for significant learning, in which the acquired knowledge is stored and can be remembered for a longer time [18].

The potential of virtual learning environments is related to the use of several elements simultaneously from mechanisms of interaction and interactivity that support the purpose of their construction, and in nursing, the use of virtual learning environments does not replace the organic relationships between subjects, the use of this technology must occur concurrently with other teaching methodologies [19].

As a strategy used to streamline and stimulate users in the navigation of virtual learning environment, videos were included, as technological tools, and by combining elements such as images, texts, and sounds in a single object, we sought to motivate and engage users in the teaching-learning process to guarantee the promotion of knowledge [20-24].

The students positively highlighted their participation in the process of building educational technologies, as well as their experience and autonomy when faced with an innovative object in their training.

For resources such as virtual environments to favor the construction of knowledge and to demonstrate knowledge gains in these educational settings, it is necessary to promote the student's interest, adapt to their needs, and adapt to their cognitive style, thus allowing a more dynamic and active learning process [25].

Conclusion

The educational technology was validated by experts, as well as being positively evaluated by students. The knowledge acquired by the students from the virtual learning environment was confirmed by statistical analyses that demonstrate the potential of this tool as an active teaching methodology to enrich and complement the traditional teaching method. The results prove the virtual learning environment to be an effective educational tool for teaching medication administration in pediatrics and neonatology and converge with the assumptions of active methodologies. Thus, the virtual learning environment of this study was established as a complementary didactic resource for theoretical teaching at the undergraduate level and the permanent education of nursing professionals.

No studies were found using design-based research in nursing with the creation of virtual learning environments, however, studies in the areas of biology, physics, medicine and psychology, education were found, which can be explained as it is a methodology that is considered to be new, although the results obtained validate the necessary efforts since the knowledge and the didactic products acquired are linked to the natural environment from which the initial problems arose and which motivated the research [26].

For future research, larger samples are suggested, for more detailed statistical power, in addition to the creation of new modules to cover some new suggested topics and the development of more videos based on the necessary demand.

Conflicts of Interest

None declared.

References

1. Wegner W. Patient safety in the care of hospitalised children: evidence for paediatric nursing. Revista Gaúcha de Enfermagem 2017 May 04;38(1):e68020. [doi: 10.1590/1983-1447.2017.01.68020]
2. Edwards IR. The WHO World Alliance for Patient Safety: a new challenge or an old one neglected? Drug Saf 2005;28(5):379-386. [doi: 10.2165/00002018-200528050-00002] [Medline: 15853440]
3. Souza MRDO, Brito EAWDS, Farias LMVC, Matias ÉO, Lima FET. Patient safety in the administration of intramuscular medication in pediatrics: assessment of the nursing practice. Rev Gaucha Enferm 2018;39(1):e2017-e2012. [doi: 10.1590/1983-1447.2018.2017-0002] [Medline: 29846477]
4. Documento de Referência do Programa Nacional de Segurança do paciente. Ministério da Saúde. Brasília; 2014. URL: http://bvsms.saude.gov.br/bvs/publicacoes/documento_referencia_programa_nacional_seguranca.pdf [accessed 2020-10-02]
5. Silva MFD, Anders JC, Rocha PK, Souza AIJD, Burciaga VB. Communication in nursing shift handover: pediatric patient safety. Texto contexto - enferm 2016;25(3):e3600015. [doi: 10.1590/0104-07072016003600015]
6. Dolmans DHJM. How theory and design-based research can mature PBL practice and research. Adv Health Sci Educ Theory Pract 2019 Dec 25;24(5):879-891 [FREE Full text] [doi: 10.1007/s10459-019-09940-2] [Medline: 31720879]
7. Seeto D, Herrington J. Design-based research and the learning designer. In: Proceedings of the 23rd Annual Conference of ASCILITE: Who's Learning? Whose Technology?. 2006 Presented at: Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education; 2006; Sydney p. 741-745.
8. Vieira GB, Cavalcante DS, Ponce de Leon CGRM, Ribeiro LM, Fonseca LMM, Menezes KR. Primeiros passos da educação permanente em enfermagem de uma unidade de terapia intensiva neonatal (First steps of nursing continuing education in neonatal intensive care unit). Revista de Enfermagem UFPE on line 2014;8(4):889-895. [doi: 10.5205/reuol.5829-50065-1-ED-1.0804201413]
9. Holanda VR. Hipermídia educacional para o ensino de doenças sexualmente transmissíveis: construção, validação e avaliação [tese]. Universidade Federal do Ceará. Fortaleza; 2014. URL: http://www.repositorio.ufc.br/handle/riufc/8310 [accessed 2018-03-05]

XSL•FO
RenderX

10.   Fehring R. The Fehring model. In: Classification of nursing diagnosis: Proceedings of the tenth conference. Philadelphia: Lippincott; 1994 Presented at: Conference on the Classification of Nursing Diagnoses; 1992; San Diego p. 55-62.

11.   Alexandre NMC, Coluci MZO. [Content validity in the development and adaptation processes of measurement instruments]. Cien Saude Colet 2011 Jul;16(7):3061-3068. [doi: 10.1590/s1413-81232011000800006] [Medline: 21808894]

12.   Polit DF, Beck CT. Fundamentos de Pesquisa em enfermagem: avaliação de evidências para as práticas da enfermagem. Porto Alegre: Artmed Editora; 2018.

13.   Worral P. Avaliação da educação para a saúde. In: Bastable SB, editor. O enfermeiro como educador. Princípios de ensino-aprendizagem para a prática de enfermagem 3 ed. Porto Alegre: Artmed Editora; 2010.

14.   Corradi MI, Silva SH, Scalabrin EE. Objetos virtuais para apoio ao processo ensino-aprendizagem do exame físico em enfermagem. Acta paul. enferm 2011;24(3):348-353. [doi: 10.1590/s0103-21002011000300007]

15.   Xelegati R, Évora YDM. Development of a virtual learning environment addressing adverse events in nursing. Rev Latino-Am Enfermagem 2011 Oct;19(5):1181-1187. [doi: 10.1590/S0104-11692011000500016]

16.   Cruz D, Santos V, Tavares V, Alencar I. Good practices of medication in intensive nursing: the experience of a training through virtual learning environment. Revista de Educação do Vale do São Francisco - REVASF 2017 Apr 4;7(12) [FREE Full text]

17.   Badiei M, Gharib M, Zolfaghari M, Mojtahedzadeh R. Comparing nurses' knowledge retention following electronic continuous education and educational booklet: a controlled trial study. Med J Islam Repub Iran 2016;30:364 [FREE Full text] [Medline: 27493908]

18.   Frota NM, Galindo Neto NM, Barros LM, Pereira FGF, Melo GAA, Caetano J. Hypermedia on peripheral venipuncture: effectiveness in teaching nursing students. Rev Bras Enferm 2018 Dec;71(6):2983-2989. [doi: 10.1590/0034-7167-2017-0205] [Medline: 30517402]

19.   Fettermann F, Alberti G, Salbego C, Kist R. Potencialidades e fragilidades dos ambientes virtuais de aprendizagem no ensino em enfermagem: revisão integrativa. Journal of Health Informatics 2017;9(4):132-136 [FREE Full text]

20.   Dalmolin A, Girardon-Perlini NMO, Coppetti LDC, Rossato GC, Gomes JS, Silva MEND. Educational video as a healthcare education resource for people with colostomy and their families. Rev Gaucha Enferm 2017 Apr 06;37(spe):e68373. [doi: 10.1590/1983-1447.2016.esp.68373] [Medline: 28403316]

21.   Stina A, Zamarioli C, Carvalho E. Effect of educational video on the student's knowledge about oral hygiene of patients undergoing chemotherapy. Escola Anna Nery - Revista de Enfermagem 2015 Apr 28;19(2):220-225. [doi: 10.5935/1414-8145.20150028]

22.   Legey AP, Maia M, Cotelli Espírito Santo A. Produção de vídeo educativo para sensibilizar profissionais de equipe de enfermagem quanto ao uso de EPIs. RECITE 2017;2(1) [FREE Full text] [doi: 10.17648/2596-058x-recite-v2n1-2]

23.   Hollands GJ, Marteau TM, Fletcher PC. Non-conscious processes in changing health-related behaviour: a conceptual analysis and framework. Health Psychology Review 2016 Feb 16;10(4):381-394 [FREE Full text] [doi: 10.1080/17437199.2015.1138093]

24.   Costa IKF, Tibúrcio MP, Melo GDSM, Leite JEL, Dantas RAN, Torres GDV. Construction and validation of a distance Basic Life Support Course. Rev Bras Enferm 2018;71(suppl 6):2698-2705. [doi: 10.1590/0034-7167-2018-0122] [Medline: 30540046]

25.   Luna I, Pinheiro P, Teixeira F. Hypermedia for teaching nursing in a digital learning environment. Brazilian Journal of Technology 2018;1(2):209-231 [FREE Full text]

26.   Kneubil F, Pietrocola M. A pesquisa baseada em design: visão geral e contribuições para o ensino de ciências (The design-based research: overview and contributions to science education). IENCI 2017 Aug 17;22(2):01-16 [FREE Full text] [doi: 10.22600/1518-8795.ienci2017v22n2p01]

XSL•FO
RenderX

<u>Original Paper</u>

Comparing Basic Life Support Serious Gaming Scores With Hands-on Training Platform Performance Scores: Pilot Simulation Study for Basic Life Support Training

Mehmet Emin Aksoy[1], MD, MSc, PhD

Acibadem Mehmet Ali Aydınlar University, Department Biomedical Device Technology, CASE (Center of Advanced Simulation and Education), Istanbul, Turkey

Corresponding Author:
Mehmet Emin Aksoy, MD, MSc, PhD
Acibadem Mehmet Ali Aydınlar University
Department Biomedical Device Technology
CASE (Center of Advanced Simulation and Education)
Kayısdagi cad. No:34
Atasehir
Istanbul, 34752
Turkey
Phone: 90 5052685158
Email: emin.aksoy@acibadem.edu.tr

Abstract

Background: Serious games enrich simulation-based health care trainings and improve knowledge, skills, and self-confidence of learners while entertaining them.

Objective: A platform which can combine performance data from a basic life support (BLS) serious game app and hands-on data based on the same scoring system is not available in the market. The aim of this study was to create such a platform and investigate whether performance evaluation of BLS trainings would be more objective compared to conventional Objective Structured Clinical Examination (OSCE) examinations if these evaluations were carried out with the platform which combines OSCE scoring criteria with sensor data retrieved from the simulator's sensors.

Methods: Participants were 25 volunteers (11 men [44.0%] and 14 [56.0] women) among Acıbadem Mehmet Ali Aydınlar University students without prior knowledge of the BLS protocol. A serious game module has been created for teaching learners the European Resuscitation Council Basic Life Support 2015 protocol. A second module called the hands-on module was designed for educators. This module includes a checklist used for BLS OSCE examinations and can retrieve sensor data such as compression depth, compression frequency, and ventilation volume from the manikin (CPR Lilly; 3B Scientific GmbH) via Bluetooth. Data retrieved from the sensors of the manikin enable educators to evaluate learners in a more objective way. Performance data retrieved from the serious gaming module have been combined with the results of the hands-on module. Data acquired from the hands-on module have also been compared with the results of conventional OSCE scores of the participants, which were obtained by watching the videos of the same trainings.

Results: Participants were considered successful in the game if they scored 80/100 or above. Overall, participants scored 80 or above in an average of 1.4 (SD 0.65) trials. The average BLS serious game score was 88.3/100 (SD 5.17) and hands-on average score was 70.7/100 (SD 17.3), whereas the OSCE average score was 84.4/100 (SD 12.9). There was no statistically significant correlation between success on trials (score ≥80/100), serious game, hands-on training app, and OSCE scores (Spearman rho test, *P*>.05). The mean BLS serious game score of the participants was 88.3/100 (SD 5.17), whereas their mean hands-on training app score was 70.7/100 (SD 17.3) and OSCE score was 84.4/100 (SD 12.9).

Conclusions: Although scoring criteria for OSCE and hands-on training app were identical, OSCE scores were 17% higher than hands-on training app scores. After analyzing the difference of scores between hands-on training app and OSCE, it has been revealed that these differences originate from scoring parameters such as compression depth, compression frequency, and ventilation volume. These data suggest that evaluation of BLS trainings would be more objective if these evaluations were carried out with the modality, which combines visual OSCE scoring criteria with sensor data retrieved from the simulator's sensors.

Trial Registration: ClinicalTrials.gov NCT04533893; https://clinicaltrials.gov/ct2/show/NCT04533893

XSL•FO

RenderX

KEYWORDS

serious gaming; medical simulation; basic life support

Introduction

The effectiveness of serious gaming has been demonstrated primarily by studies of higher education, government, corporate, and military environments. Serious gaming and medical simulation techniques, which are 2 complementary modalities for training of tomorrow's physicians, have gained popularity in the past decade. Olszewski and Wolbrink [1] emphasize that serious games are increasingly being used for medical education. Blended learning methodology combines serious gaming with medical simulation programs and enhances the efficiency and effectiveness of training programs [2]. E-learning by using serious gaming refers to the use of internet technologies to deliver an interactive educational content that enhances knowledge and performance [3,4]. Prensky [5] concluded that game technologies can produce a great deal of learning and positive effects such as quickly mastering and applying new skills and information, thinking strategically, and persisting to solve difficult problems. David and Michael [6] described serious game methodology as using game-based simulations that do not have entertainment, enjoyment, or fun as their primary purpose.

Nowadays, after providing knowledge using serious game–based apps, simulation-based trainings and Objective Structured Clinical Examination (OSCE) are used to assess hands-on training performance of trainees [7,8].

Basic life support (BLS) is the most frequently organized training course for both medical and nonmedical trainees. Depending on the country, BLS certifications have to be renewed on a yearly basis. Smith et al [9] inspected the decline of BLS skills retention in 3, 6, 9, and 12 months. In BLS, even after 3 months, test scores of trainees decreased by 33.3% compared to their posttraining test scores. Decrease in knowledge and skills are more dramatic after 6, 9, and 12 months [9]. Therefore, BLS providers need more frequent refresher training to improve and maintain their knowledge, skills, and self-confidence.

Depending on these facts, a multiplatform-compatible interactive serious game module was created in our simulation center to teach adult BLS Protocol based on the European Resuscitation Council (ERC) 2015 Criteria [10]. As serious gaming performance analysis is only used for knowledge domain assessment, effectivity of BLS trainings also has to be assessed by observing hands-on performance of learners. Therefore, it has been decided to develop a tablet-based app for educators, which is capable of retrieving sensor data such as compression depth, compression frequency, ventilation volume from the manikin via Bluetooth and includes OSCE criteria for BLS training. Sensor data retrieved from the manikin enable educators to evaluate learners in a more objective way. Both performance data from the serious gaming module and hands-on training app can be stored on the same learning management system (LMS), which allows educators to track and compare performances on individual basis. There is no available platform in the market which can combine performance data from the BLS serious game app and hands-on data based on the same scoring system.

As OSCEs are mainly based on observation of the simulation-based trainings, this new platform, which combines visual OSCE scoring criteria with sensor data retrieved from the simulator's sensors, would be able to provide more objective and precise scores for BLS trainings. The hypothesis of this study is that whether performance evaluation of BLS trainings would be more objective compared to conventional OSCEs, if these evaluations were carried out with the modality which combines visual OSCE scoring criteria with sensor data retrieved from the simulator's sensors.

Methods

Recruitment

After obtaining approval from the Acıbadem Mehmet Ali Aydınlar University Ethics Committee, participants were chosen among volunteers and asked to fill and sign consent forms. The participants of this study included 25 volunteers among Acıbadem Mehmet Ali Aydınlar University students without prior knowledge of BLS protocol; 11 participants (44%) were male and 14 (56%) were female. Volunteers with prior BLS training were excluded from the study.

Study Design

After a brief introduction of the serious gaming module, the participants were asked to log into the system with their own usernames and passwords. After completing the training mode of the serious game module, they were asked to choose the self-test mode of the module. The participants' number of attempts was not limited. Because of the COVID-19 pandemic, the participants were asked to practice their hands-on skills in the simulation center under the supervision of educators on the same day in order to minimize the risk of contamination. The new hands-on training app, which combines OSCE scoring criteria with sensor data retrieved from the simulator's sensors, was used for hands-on performance evaluation. After familiarization with the system using the self-training mode, the participants were asked to use the BLS hands-on training app with the simulator under the supervision of the educator. The simulation sessions were also recorded in order to use these recordings for conventional OSCE performance evaluation. Conventional OSCE scores of participants were obtained by watching the recorded sessions of BLS trainings. The study design is shown in Figure 1.

XSL•FO

RenderX

Figure 1. Design of the study flow.

Learning Management System Data Collection

Serious Game Self-Study Mode → Serious Game Exam Mode → Hands-on App Familiarization → Hands-on Training Mode → Hands-on App Exam Mode

Video Recording for OSCE
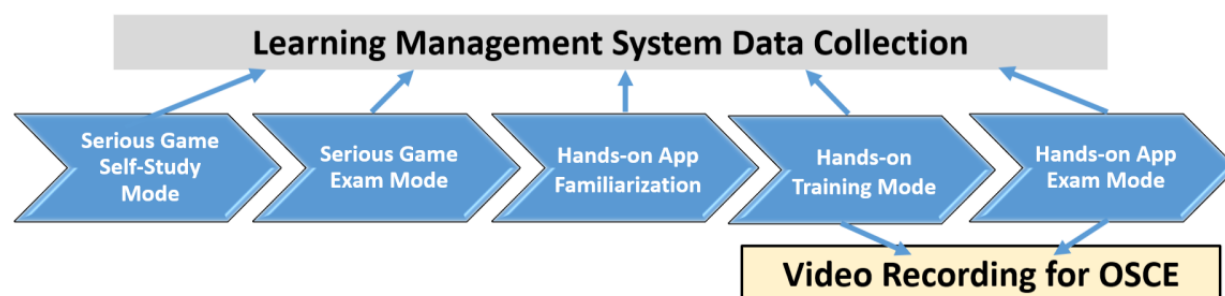
Serious Game–Based BLS Training App

The tablet-based serious game module is compatible with 2015 Basic Life Support Algorithm of the ERC and includes an LMS, a learning record store (LRS), and a 3D visualization engine (VE) [11,12]. LMS and LRS enable to store users' credentials in a shared database. Through the interaction between LRS and VE every action performed by the user can be followed on the VE side and experience API (xAPI) calls corresponding to each action can be made to the LRS [12,13]. During serious game play, whenever the user has entered a predefined area or whenever the user has triggered some kind of predefined interaction with a significant object that is part of the virtual environment, an xAPI event is automatically generated [12-14]. In order to provide this interaction, a software library has been developed as a unity extension. The main function of this library is to perform the necessary xAPI web service calls, to the LRS servers, over the HTTP protocol. All actions, which may need to be performed by the user as part of an education scenario, are defined in terms of the actor, verb, and object basic parameters. As these actions are performed by the user, the library automatically issues the xAPI calls corresponding to each action, and the LRS keeps track of each significant action through these calls [12-14]. In addition to actions directly performed by the user, it is also possible to collect xAPI calls automatically for actions that have not been performed within a given time limit. The library also implements the security and user authentication features necessary for keeping proper

records. For this purpose, the basic access authentication method which is part of the HTTP protocol standard is used. The serious game study was named as "3D SIM Training Basic Life Support." 3D SIM Training Basic Life Support is a multilingual serious game app with 3D and interactive features, based on ERC 2015 Guidelines. The app consists of a training mode, a self-test mode, and a review mode enabling users to track their previous records. The training module guides the user step by step through a scenario. The user has to go through a training mode followed by the self-test mode. The user is not in the passive learner mode but has an active role in the game. The user is first instructed about the correct actions to be taken and then interacts with the 3D software while playing the rescuer role interactively. All main steps of the ERC 2015 Adult BLS algorithm are covered in the BLS Serious Game app. The user is expected to finish 1 stage correctly before continuing with the next stage. The app is developed as a 3D real-time game enabling the user to interact and view the scene from different angles. A 3D animation-like chest compression timing clarifies some of the steps within the BLS algorithm and gives the trainees the feeling of a chest compression frequency between 100 and 120 compressions/minute. Screen captures from the Adult BLS Module are presented in Figure 2. The animation of blood flow during CPR informs the users about the aim of the chest compression. The right locations for checking carotid pulse, chest compressions, and automated external defibrillator pad locations are also indicated during the self-training mode.

Figure 2. Screen captures from the Basic Life Support Serious Game Module.

Once the user successfully finishes the training module, he/she has to complete the self-test module. Unlike the training module, users are expected to perform the correct steps with correct timing without hints or instructions to follow in the self-training mode. The SCORM (Sharable Content Object Reference Model)-xAPI standard has been used for this serious game

module due to its advantages such as supporting a variety of content types, simplicity for implementing, and supporting offline or disconnected scenarios. Because of these advantages, the SCORM-xAPI standard has been chosen as the methodology to track the digital learning content [4,15]. Because of its advanced features, x-API can be considered as a replacement

for previous e-learning standards. In this serious game module, all actions and their timings are tracked using the SCORM-xAPI standard and the outcome of the "game" depends on the user's success. At the end of the scenario, the user is presented with an easily understandable performance chart. The user's performance can also be reported to an internet-based server to create test results for offline evaluation. The reported performance summary incorporates all the expected steps and also shows if they have been performed in the right order and correct timing as shown in Figure 3.

Figure 3. Performance report of BLS serious game app.
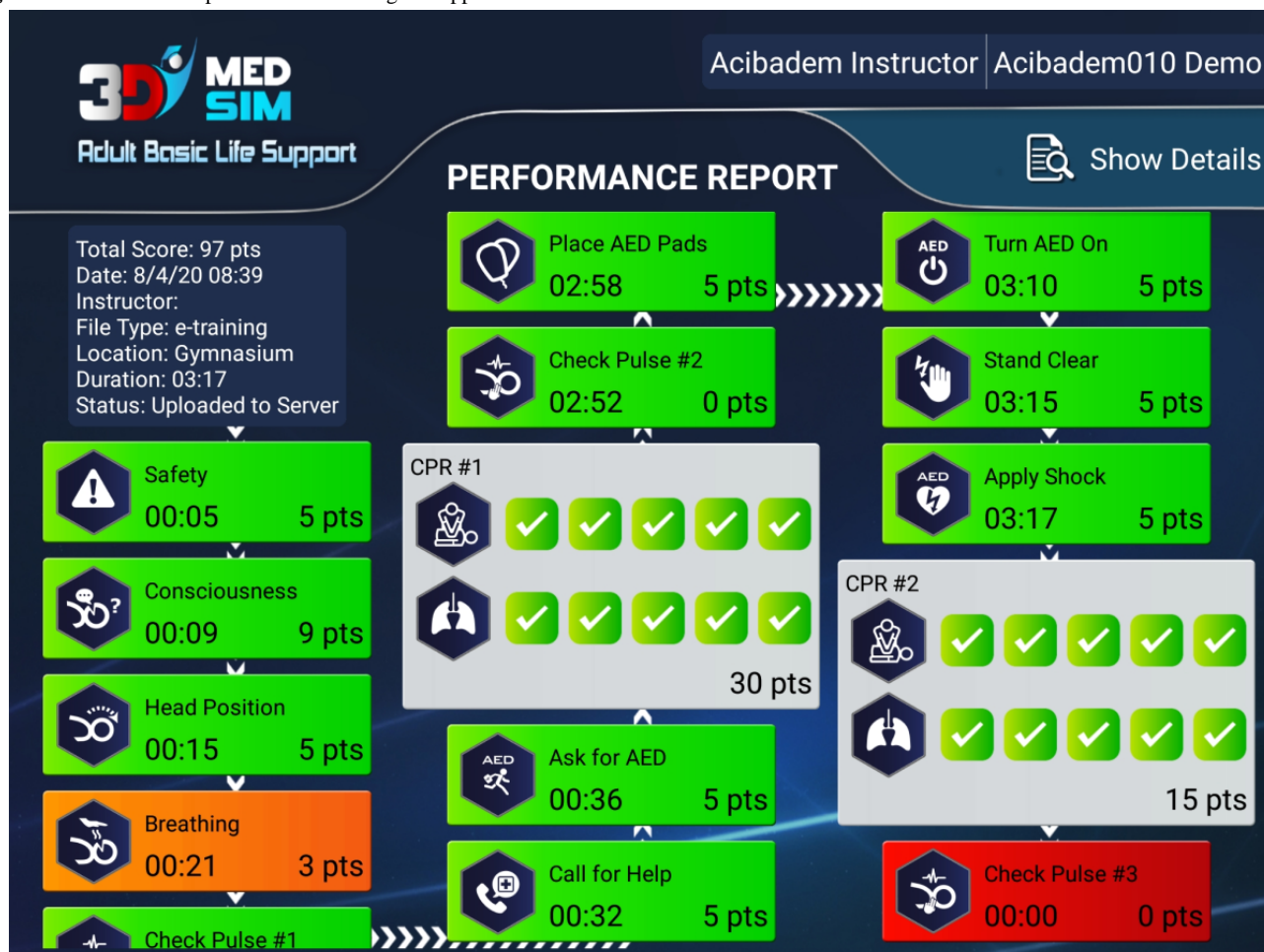
Each step of the user's performance is evaluated with 3 different colors. Green symbolizes that the step has been successfully applied in appropriate time. Orange symbolizes that the step has been successfully applied but timing was wrong. Red symbolizes that the step has not been applied at all. A virtual 3D simulation environment is developed for training of medical staff for various medical processes. Integration and coordination of 3 main parts, including 3D VE, LMS, and LRS, were necessary for the functionality of the whole system. Technical suitability, wide platform support (iOS, Android, Windows, Linux, and web), large user database, effective support, and favorable licensing conditions are the parameters for selection of a 3D VE in an educational simulation project [16-19]. Among the criteria mentioned above, technical suitability deserves a more detailed explanation and covers all possible technical requirements specific to the implementation of a 3D education and simulation app integrated with an LMS. This consideration arises because modern VEs are not strictly limited to visualization anymore. Essentially, they are contained within larger frameworks which provide the functionality required for developing entire apps [16-19]. In general, currently available engines have been designed, first and foremost, with the needs of game developers in mind [20]. For realization of this study, performances of a number of 3D visualization and game engines have been evaluated and compared. After evaluating available game engine products, the Unity game engine has been found the most suitable for this study, because it covers all of the criteria outlined above [21]. Owing to the highly extensible architecture of the Unity engine, the users greatly extend the capabilities of the engine by creating a large number of useful and affordable extensions [21].

Learning Management System (LMS)

LMS is a software designed for managing user learning interventions and delivering learning content to the learners [13,14,22]. LMS is the key for flexible, interactive, multimedia, and decentralized teaching [12-14]. The advantages of using LMS for the educator and learner has been revealed in several studies [23,24]. Therefore, a custom-designed LMS has been integrated to the serious game module in this project.

Learning Record Store

LRS is a cloud-based service that deals with learning information storage and retrieval of learning information [13,25]. For the LRS, the open source reference LRS implementation, provided by advanced distributed learning, has been selected for this project [26].

Scoring System

A successful serious game project requires the LMS, the LRS, and the 3D VE to operate in harmony. For this purpose, a software library has been developed as a Unity extension. The main function of this library is to record the track of every action during the game through xAPI calls [12,13,27]. In addition to

the actions directly performed by the user, it is also possible to have xAPI calls issued automatically for actions that have not been performed within a given time limit. The library also implements the security and user authentication features necessary for keeping proper records. For this purpose, the basic access authentication method which is part of the HTTP standard was used. Login usernames and passwords were sent to the participants by email after they have downloaded the serious game from iTunes or Google Play. The scores of the participants were stored with the help of the LMS. The participants' scores were downloaded from the LMS of the serious game module and converted to an MS Excel file format. The same scoring criteria were used for BLS serious gaming app, hands-on app, and OSCE performance analysis as shown in Table 1.

Table 1. Score comparisons according to gender.

Condition	Scoring		
	BLS[a] serious game	Hands-on app with manikin	OSCE[b]
Check Safety	5	5	5
Check Consciousness	9	9	9
Head Tilt	5	5	5
Check Breathing	6	6	6
Call for Help	5	5	5
Get AED[c]	5	5	5
Check Carotid Pulse	0	0	0
1st CPR30[d]	4	4	4
1st Ventilation	2	2	2
2nd CPR30	4	4	4
2nd Ventilation	2	2	2
3rd CPR30	4	4	4
3rd Ventilation	2	2	2
4th CPR30	4	4	4
4th Ventilation	2	2	2
5th CPR30	4	4	4
5th Ventilation	2	2	2
2nd Carotid Pulse	0	0	0
AED Use Pad	5	5	5
AED On Off	5	5	5
Stand Clear	5	5	5
AED Shock	5	5	5
Final CPR	15	15	15
Check Rhythm	0	0	0
Total score	100	100	100

[a]BLS: basic life support.

[b]OSCE: Objective Structured Clinical Examination.

[c]AED: automated external defibrillator.

[d]CPR: cardiopulmonary resuscitation.

Hands-On Module

Hands-on module is a platform developed for BLS training featuring hands-on training. The system is built around an Android app, which communicates with a commercially available BLS manikin (CPR Lilly; 3B Scientific GmbH). The system allows learners to practice, to track their progress using cloud services, and allows educators to conduct BLS examinations with OSCE criteria combined with sensor data

obtained from the manikin via Bluetooth. Data acquired from the hands-on module are also stored by the same LMS, which is also used by the BLS serious gaming app. In this way, results from the BLS serious gaming app and hands-on module are stored at individual level by the LMS, so that each learner's performance data at the knowledge level and hands-on skill level can be tracked by the educator. The system architecture is shown in Figure 4. The system comprises 5 parts as shown in Figure 4.

Figure 4. System architecture of the platform.
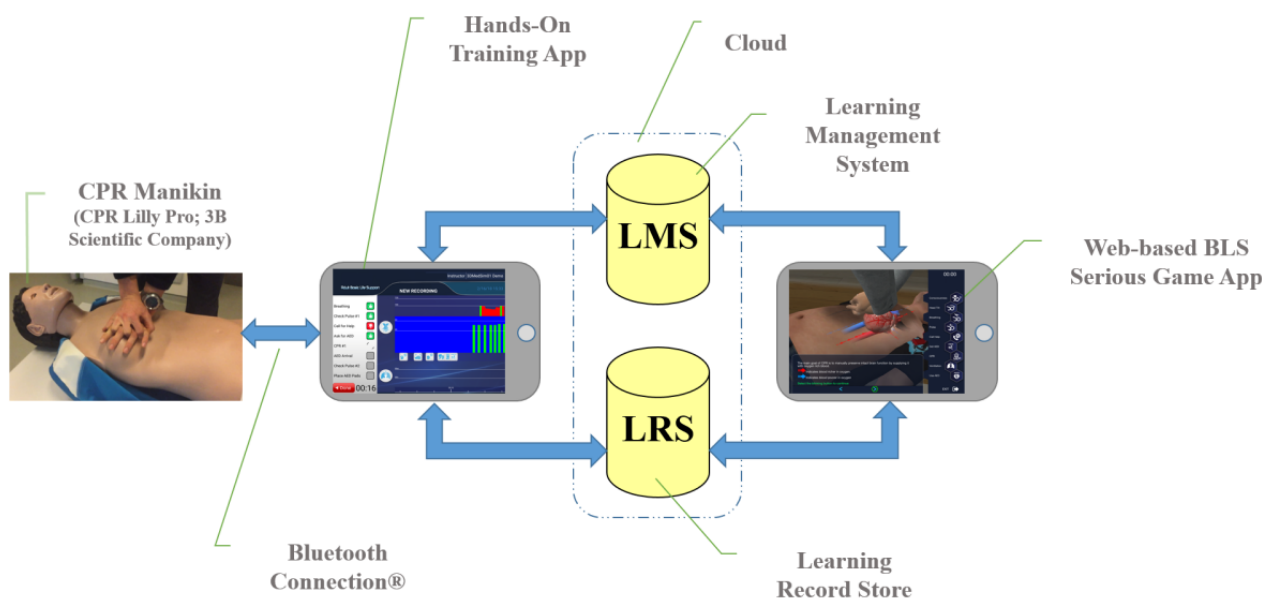
The LRS stores the progress of the learners' hands-on training sessions in a detailed step-by-step fashion. This information is used to provide the learners with a history of past sessions as well as an up-to-date report card. LMS also accesses the LRS to generate reports for instructors. The hands-on BLS training platform supports any xAPI-compliant third-party LRS. xAPI is an e-learning specification that makes it possible to collect data within online and offline training activities. It is a shared format for receiving and sending learning data between multiple systems. The app provides the learners and instructors with a visual interface to access the past records and current report card. It also communicates with the manikin to record BLS sessions. The manikin contains internal sensors, which provide measurements for chest compression depth and ventilation volume. These measurements can be provided to Bluetooth clients in real time. An instructor must first login to the app to use the system. The app validates the instructor's credentials

against the LMS. Learners must use their credentials to allow the instructor access their data. The app confirms with LMS that the 2 users are entitled to each other for the BLS training course. A screen capture of the hands-on app is shown on Figure 5. The hands-on app combines visual checks based on BLS OSCE criteria and live sensor values for chest compression depth and frequency and ventilation volume obtained from the BLS manikin's sensors via Bluetooth.

The app obtains LRS access parameters and uses them to download records of past BLS sessions of the learners, along with their scores for these sessions and previous BLS serious game scores as shown in Figure 6.

Choosing to review a past session displays a report card for the session, which contains a score and the BLS process broken down into steps. Each step marked with a score component, timing, evaluation, and rationale as shown in Figure 7.

Figure 5. Screen capture of the Hands-on app.
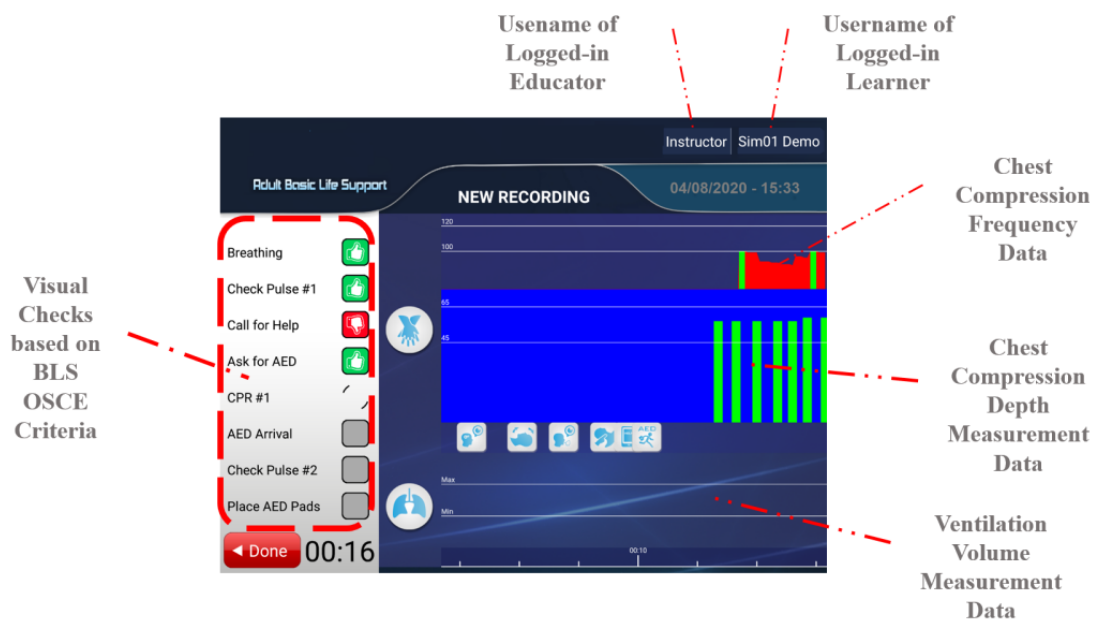
Figure 6. Hands-on app past performance screen.

XSL•FO

RenderX
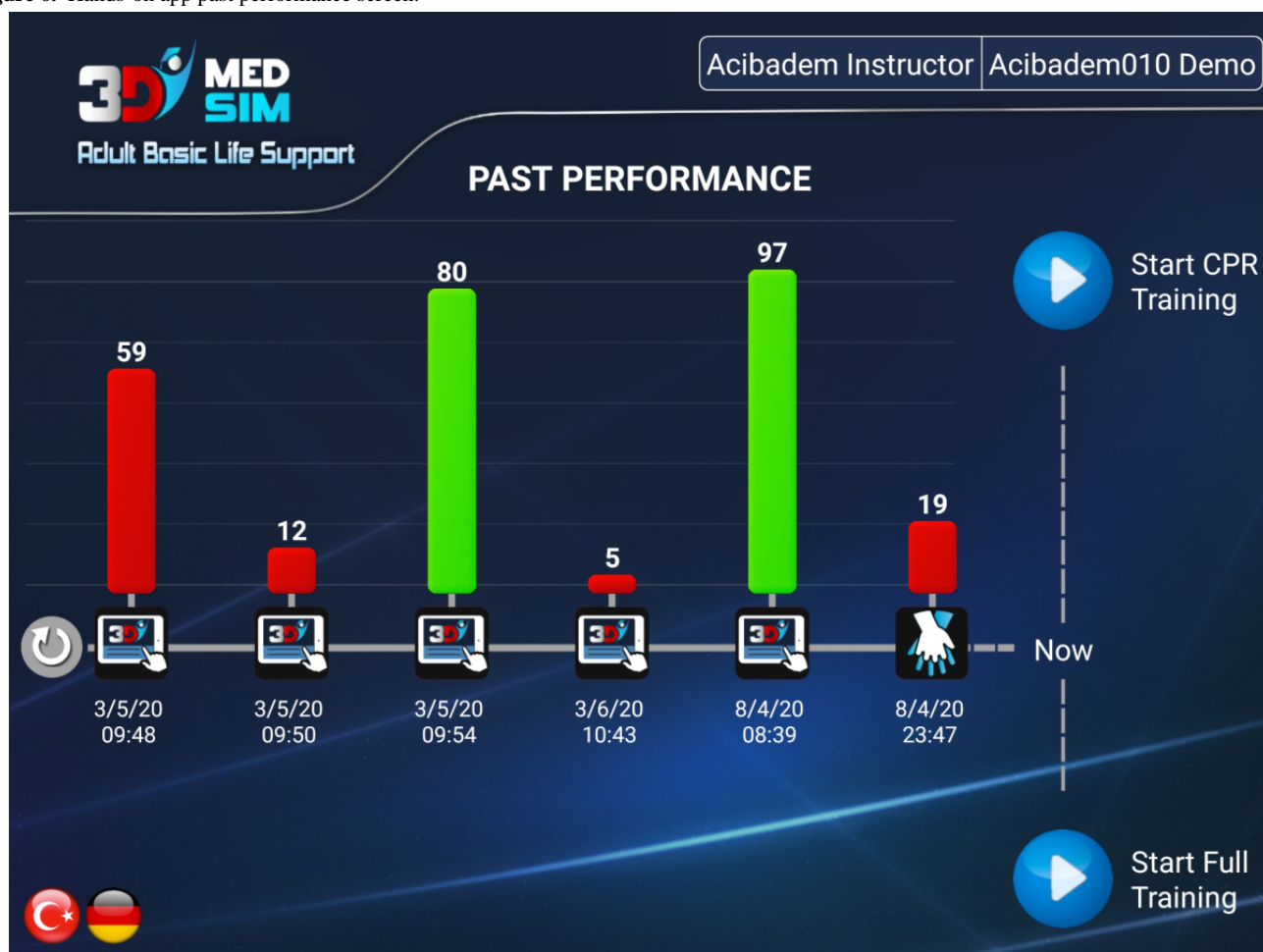
Figure 7. Performance report screen hands-on app.
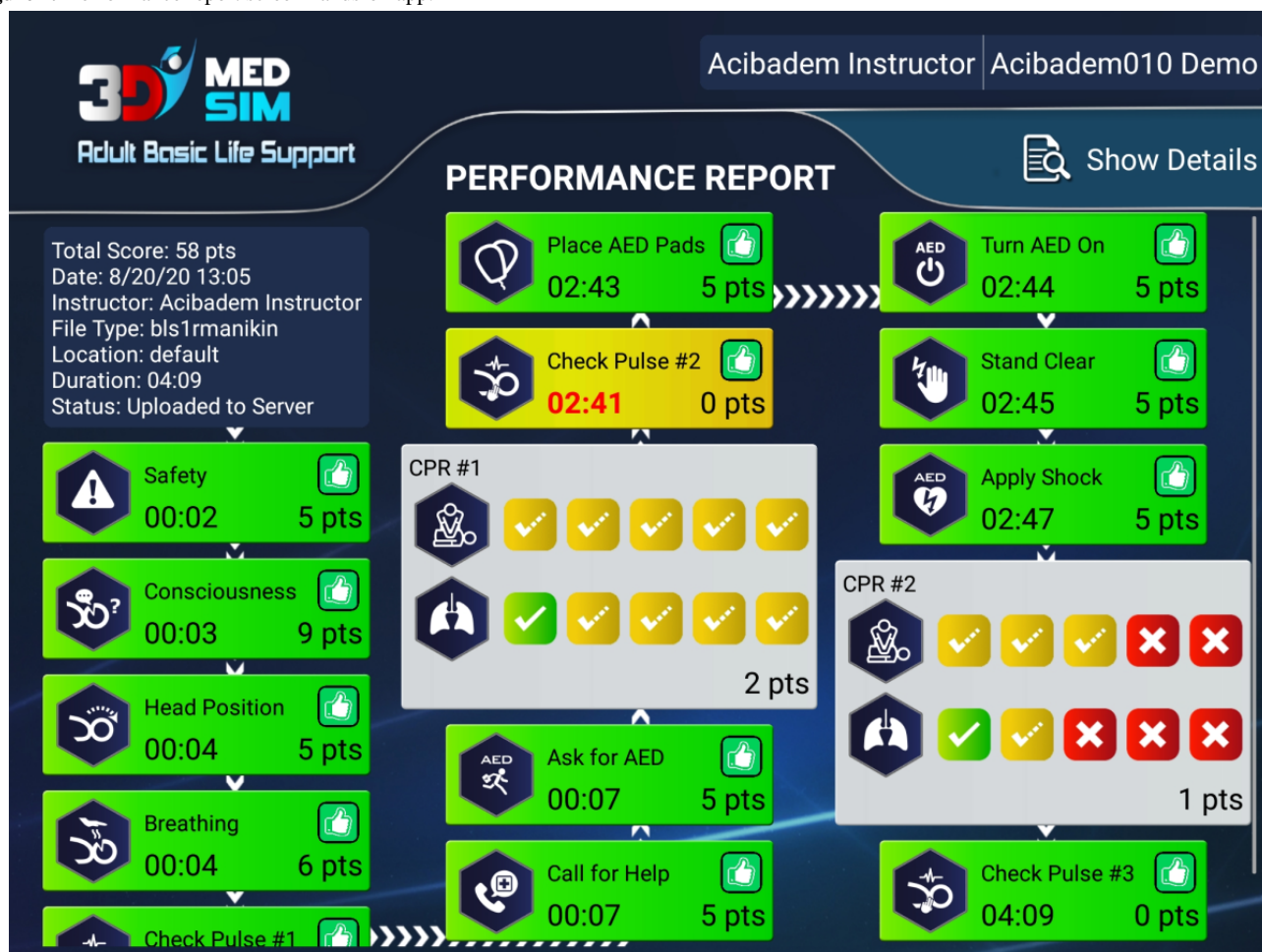
When reviewing a past session, it is also possible to see a detailed timeline of all events in that session, including synchronized BLS actions as shown in Figure 7. Starting a new session triggers the manikin discovery-auto-connect procedure and is possible in 2 different modes. In the "BLS Training" mode, only BLS actions are allowed and a metronome is provided for assistance. The session is not recorded. In the "Full Training" mode, all BLS events are marked along with the BLS algorithm. This mode is intended to be used by the instructor who can mark events that are not electronically tracked to augment the data received from the manikin with their own timestamps. In this mode, the session may be scored and permanently recorded. The app accesses the manikin wirelessly through a secure Bluetooth interface—the BLS Manikin (CPR Lilly; 3B Scientific GmbH). Therefore, the Android device must first be paired with the manikin using the app setup wizard. This wizard also ensures that the app knows which manikin to receive data from. When opening a session, the app creates a Bluetooth connection to the manikin using the serial port profile model. This connection is bidirectional and allows the app to control various functions of the manikin (such as eyes or simulated heartbeat) while also receiving data from it. All of this communication is delegated to a manikin service that the app installs to the system (Figure 8).

Figure 8. Android App Architecture.
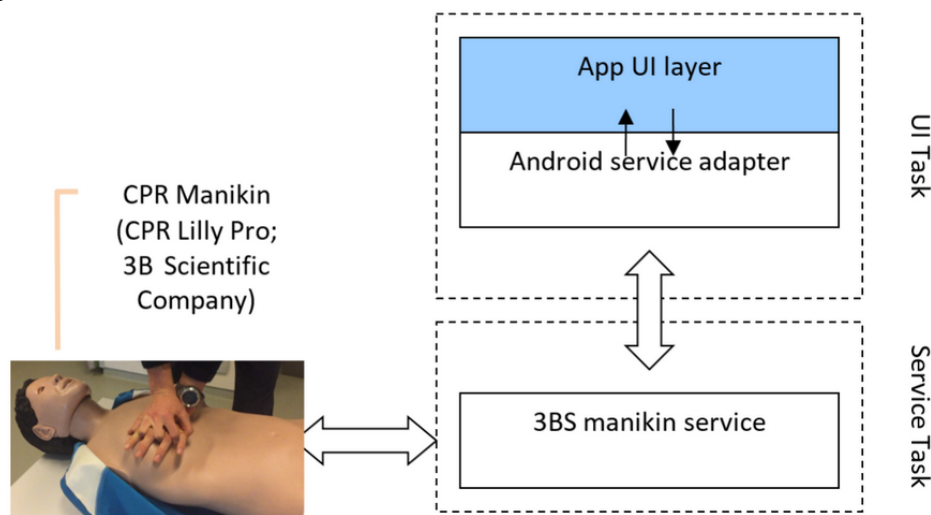
Before recording, the app first clears the internal buffers of the manikin service for a new session and then signals the receipt of data. During the session, a steady stream of sensor readings is sent by the manikin to the manikin service. These readings are tracked linearly in order to detect rising and falling edges, which are used to deduce compression and ventilation events. These events are collected into data buffers, which are then transmitted to the Android app. The Android app filters the generated events for possible misdetections and augments them with additional sensor data present in the stream. Such sensor data are received momentarily, but linearly averaged by the app over a window around the event times. The detected and augmented BLS events are then added to the session along with their timestamps and evaluated when the session is over.

Statistical Analysis

The normality of continuous variables was investigated by Shapiro–Wilk test. Descriptive statistics were presented using mean and standard deviation, median, and range. For comparison of 2 non-normally distributed groups Mann–Whitney U test was used. Spearman rho correlation test was used for the correlation between 2 continuous variables that are independent and normally distributed. Statistical significance was accepted when 2-sided P value was lower than .05. Statistical analysis was performed using the MedCalc Statistical Software version 12.7.7 (MedCalc Software Ltd) [28].

Results

To be successful from the serious gaming module, participants had to score 80/100 or above. Participants got 80/100 and above in an average of 1.4 (SD 0.6) trials.

The average BLS serious game score was 88.3/100 (SD 5.1), hands-on average score was 70.7/100 (SD 17.3), and the OSCE average score was 84.4/100 (SD 12.9), as shown in Table 2.

Serious game scores and number of trials of 25 participants are shown in Figure 9.

Table 2. Gender distribution and BLS[a] serious game, hands-on, and OSCE[b] scores.

Gender and scores in modules	Values
Gender	
Male, n (%)	11 (44)
Female, n (%)	14 (56)
Success on which trial (80 and above)	
1, n (%)	17 (68)
2, n (%)	6 (24)
3, n (%)	2 (8)
BLS Serious Game first trial	
Mean (SD)	81 (16.9)
Median (range)	86 (22-100)
BLS Serious Game second trial	
Mean (SD)	80 (13.2)
Median (range)	83.5 (55-93)
BLS Serious Game third trial	
Mean (SD)	84 (2.8)
Median (range)	84 (82-86)
Success on the serious game trial (80 or above was required for success)	
Mean (SD)	1.4 (0.6)
Median (range)	1 (1-3)
BLS Serious Game best score	
Mean (SD)	88.3 (5.1)
Median (range)	89 (80-100)
Hands-on app	
Mean (SD)	70.7 (17.3)
Median (range)	73 (37-95)
OSCE	
Mean (SD)	84.4 (12.9)
Median (range)	89 (55-100)

[a]BLS: basic life support.

[b]OSCE: Objective Structured Clinical Examination.

XSL•FO
RenderX

Figure 9. Serious game scores und number of trials of 25 participants.

The score for being successful from the serious gaming module is 80/100 or above. Participants scored 80/100 and above in an average of 1.4 (SD 0.6) trials.
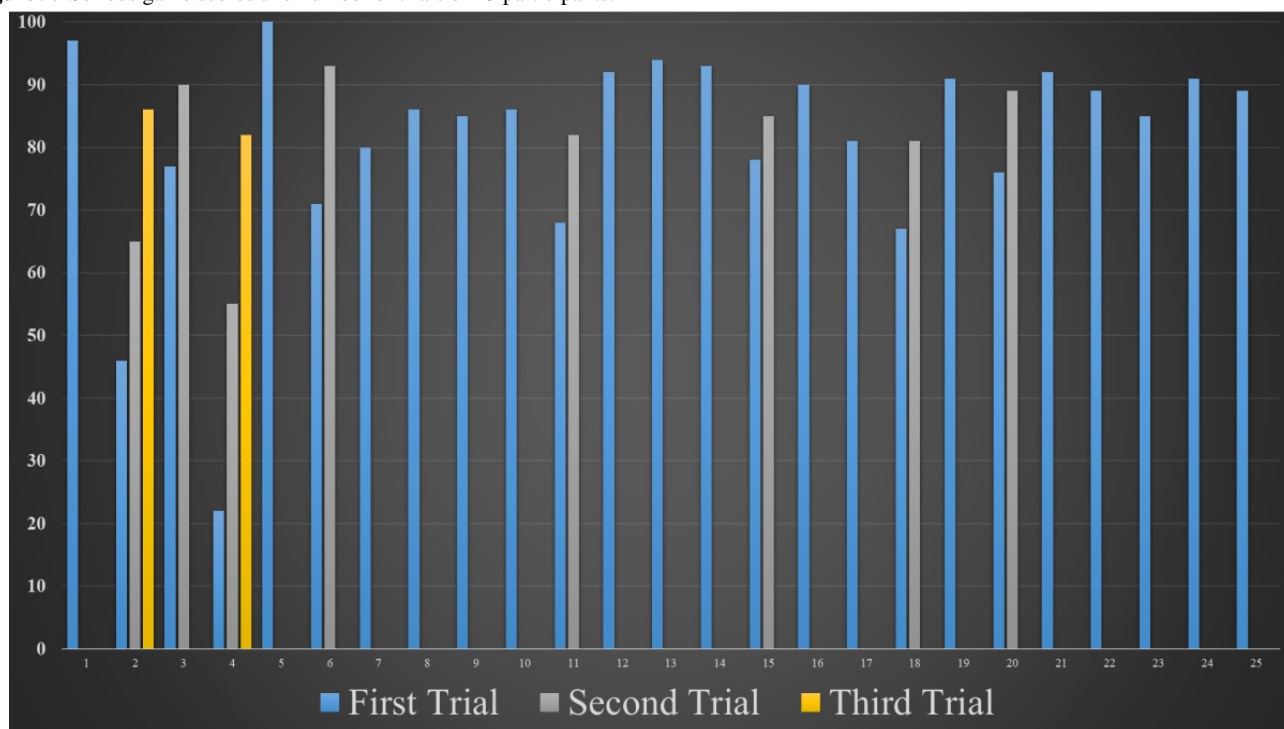
The average BLS serious game score was 88.3/100 (SD 5.1), hands-on average score was 70.7/100 (SD 17.3), and the OSCE average score was 84.4/100 (SD 12.9) as shown in Table 3.

Table 3. Comparisons according to gender[a].

Score comparison	Male	Female	P-value
Success on trial (80 and above)			.936
Mean (SD)	1.3 (0.5)	1.4 (0.7)	
Median (range)	1 (1-2)	1 (1-3)	
Tablet			.267
Mean (SD)	89.8 (6)	87.2 (4.2)	
Median (range)	90 (81-100)	87.5 (80-93)	
Hands-on			.166
Mean (SD)	76.2 (17.1)	66.4 (16.9)	
Median (range)	77 (37-95)	70 (37-90)	
OSCE[b]			.166
Mean (SD)	81.2 (12.3)	86.8 (13.3)	
Median (range)	84 (55-100)	90 (55-100)	

[a]By Mann–Whitney U test.

[b]OSCE: Objective Structured Clinical Examination.

According to gender, there was no statistically significant difference in terms of success on any of the trials: tablet, hands-on, and OSCE distributions (Mann–Whitney U test, P>.05) as seen in Table 3.

As shown in Table 4, there was no statistically significant correlation between success on any trial (score ≥80), serious game, hands-on training app, and OSCE (Spearman rho test, P>.05). Serious game, hands-on training app, and OSCE scores of 25 participants are shown in Figure 10.

Table 4. Correlation analysis of serious game, hands-on training app, and OSCE scores[a].

Correlation	Success on which trial (score ≥80)	Tablet	Hands-on	OSCE[b]
Success on which trial (score ≥80)				
r	1.000	–0.319	–0.196	0.050
P		.120	.347	.811
Tablet				
r	–0.319	1.000	0.319	0.095
P	.120		.120	.651
Hands-on				
r	–0.196	0.319	1.000	–0.052
P	.347	.120		.806
OSCE				
r	0.050	0.095	–0.052	1.000
P	.811	.651	.806	

[a]By Spearman rho test.

[b]OSCE: Objective Structured Clinical Examination.

Figure 10. Serious game, hands-on training app and OSCE scores of 25 participants.
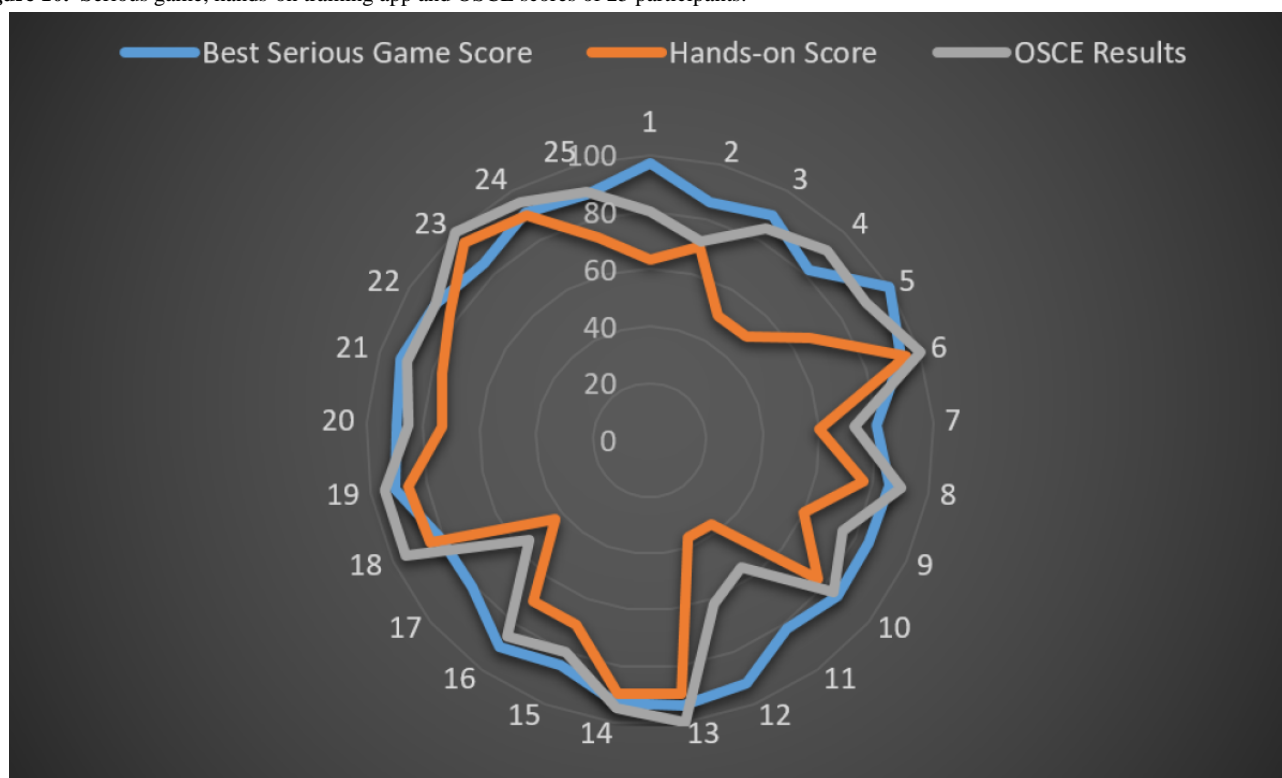
Discussion

The advantages of using serious gaming versus classroom lectures for BLS trainings as hands-on course preparation tool have been demonstrated in several studies [29-31]. There is no commercially available platform combining serious game performance data with hands-on training data obtained by visual checks based on OSCE criteria and sensor data such as compression depth, compression frequency, and ventilation volume from a BLS manikin via Bluetooth before this study.

By storing these data of trainees at an individual level in the same LMS, educators were able to see the progress of the trainings and had precise data about compression depth, compression frequency, and ventilation volume retrieved from the manikin's sensors. In conventional OSCE-based BLS training examinations, educators have to estimate compression depth, compression frequency, and ventilation volume depending on their visual estimation. By using the new platform combining BLS serious game scores and hands-on training data obtained with visual OSCE criteria and sensor data, assessment

of the knowledge domain and practical domain could be carried out with performance data that can be retrieved from the LMS at an individual level. Statistical analysis of this study revealed that there was no statistically significant difference in terms of success in any of the trials, (serious game, hands-on, and OSCE) and their score distributions (Mann–Whitney U test, $P>.05$) according to gender as shown in Table 3. All participants were given unlimited access for using the serious game module and all of them could reach 80/100 scores or above in an average of 1.4 (SD 0.6) trials. These data revealed that the knowledge domain could be assessed with the examination mode of the serious gaming module. By contrast, no statistically significant correlation was detected between the participants' serious game, hands-on training app, and OSCE scores (Spearman rho test, $P>.05$) as shown in Table 4. The mean BLS serious game score of the participants was 88.3 (SD 5.1), whereas their mean hands-on training app score was 70.7 (SD 17.3) and OSCE score was 84.4 (12.9). Although identical scoring criteria were used for hands-on training app and OSCE, OSCE scores were 17% higher than hands-on training app scores. After analyzing the difference in scores between the hands-on training app and OSCE, it was found out that these differences originated from scoring parameters such as compression depth, compression frequency, and ventilation volume. These data support the hypothesis of this study that evaluation of BLS trainings would be more objective if these evaluations were carried out with the modality, which combines OSCE scoring criteria with sensor data retrieved from the simulator's sensors.

Limitations

Only 1 of the 25 participants (4%) experienced login issues during authorization and due to internet connectivity issues. Two software bugs were encountered during the implementation phase and were all solved in the third version of the serious gaming software. The other problem was a Bluetooth connectivity issue between the hands-on app and BLS manikin in 4 of the 25 (16%) participants. This problem was resolved by removing other Bluetooth devices from the training area. The limitation of this study was that the number of participants was limited to 25. This was due to the COVID-19 pandemic, as the number of students should be limited and diluted in order to minimize any contamination risk.

Conclusions

The advantages of using serious games before simulation sessions has been proven by several studies. This study revealed that using the platform combining serious game performance data with hands-on training data obtained by combining visual OSCE criteria and sensor data such as compression depth, compression frequency, and ventilation volume from the BLS manikin via Bluetooth enables educators to have an overview of learners' serious game and hands-on performance data. The platform provides more accurate data in comparison with visual observation when assessing compression depth, compression frequency, and ventilation volume compared with conventional OSCE examinations. This platform is planned to be used with more learners and educators in the upcoming studies. A new augmented virtuality–based module, which will provide mixed reality–based BLS training, is being developed at our center. This new module will be replacing the hands-on training app in our upcoming studies. Using this module, tracking the manikin in virtual space, having haptic response from the manikin, and retrieving its sensor data will be possible. Similar studies are planned by using the mixed reality–based platform in the upcoming months, when the system is ready for use.

Conflicts of Interest

None declared.

References

1. Olszewski AE, Wolbrink TA. Serious Gaming in Medical Education: A Proposed Structured Framework for Game Development. Simul Healthc 2017 Aug;12(4):240-253. [doi: 10.1097/SIH.0000000000000212] [Medline: 28027076]
2. Thorne K. Blended learning how to integrate online & traditional learning. London: Kogan Page; 2003:35-70.
3. Ruiz JG, Mintzer MJ, Leipzig RM. The Impact of E-learning in Medical Education. Acad Med 2006 Mar;81(3):207-212. [doi: 10.1097/00001888-200603000-00002] [Medline: 16501260]
4. Sanderson PE. E-Learning: strategies for delivering knowledge in the digital age. The Internet and Higher Education 2002 Jun;5(2):185-188. [doi: 10.1016/s1096-7516(02)00082-9]
5. Prensky M. Digital game-based learning. Comput Entertain 2003 Oct;1(1):21. [doi: 10.1145/950566.950596]
6. Michael D, Chen S. Serious Games: Games That Educate, Train and Inform. Canada: Thomson Course Technology; 2006:17-43.
7. Sturpe DA, Huynh D, Haines ST. Scoring objective structured clinical examinations using video monitors or video recordings. Am J Pharm Educ 2010 Apr 12;74(3):44 [FREE Full text] [doi: 10.5688/aj740344] [Medline: 20498737]
8. Kelly MA, Mitchell ML, Henderson A, Jeffrey CA, Groves M, Nulty DD, et al. OSCE best practice guidelines-applicability for nursing simulations. Adv Simul (Lond) 2016 Apr 2;1(1):10 [FREE Full text] [doi: 10.1186/s41077-016-0014-1] [Medline: 29449979]

XSL•FO

RenderX

9. Smith KK, Gilcreast D, Pierce K. Evaluation of staff's retention of ACLS and BLS skills. Resuscitation 2008 Jul;78(1):59-65. [doi: 10.1016/j.resuscitation.2008.02.007] [Medline: 18406037]

10. Perkins GD, Handley AJ, Koster RW, Castrén M, Smyth MA, Olasveengen T, Adult basic life support and automated external defibrillation Collaborators. European Resuscitation Council Guidelines for Resuscitation 2015: Section 2. Adult basic life support and automated external defibrillation. Resuscitation 2015 Oct;95:81-99. [doi: 10.1016/j.resuscitation.2015.07.015] [Medline: 26477420]

11. Koster RW, Baubin MA, Bossaert LL, Caballero A, Cassan P, Castrén M, et al. European Resuscitation Council Guidelines for Resuscitation 2010 Section 2. Adult basic life support and use of automated external defibrillators. Resuscitation 2010 Oct;81(10):1277-1292 [FREE Full text] [doi: 10.1016/j.resuscitation.2010.08.009] [Medline: 20956051]

12. Ocheja P, Flanagan B, Ogata H. Connecting decentralized learning records: a blockchain based learning analytics platform. In: LAK '18: Proceedings of the 8th International Conference on Learning Analytics and Knowledge. New York: Association for Computing Machinery; 2018 Presented at: 8th International Conference on Learning Analytics and Knowledge; 5-9 March 2018; Sydney, New South Wales, Australia p. 265-269 URL: https://dl.acm.org/doi/proceedings/10.1145/3170358 [doi: 10.1145/3170358.3170365]

13. Clemnte FJG, Cubillo LT, Dormido S, Salzmann C, Gillet D. Collecting experience data from remotely hosted learning applications. In: Auer ME, Zutin DG, editors. Online Engineering & Internet of Things. Cham, Switzerland: Springer International Publishing; Jan 2018:170-181.

14. Muloz MLA, Orfali CH. The Role of Learning Management Systems (LMS) in Educational Environments: Didactic and Technical Use. In: ICERI2018 Proceedings. Valencia, Spain: IATED; 2018 Presented at: 11th Annual International Conference of Education, Research and Innovation; 12-14 November, 2018; Seville, Spain p. 10473-10479. [doi: 10.21125/iceri.2018.0976]

15. Jin Q. In: Jin Q, editor. Distance Education Environments and Emerging Software Systems New Technologies. Hershey, PA: IGI Global; May 31, 2011:3-342.

16. Bijl JL, Boer CA. Advanced 3D visualization for simulation using game technology. In: Proceedings of the 2011 Winter Simulation Conference (WSC). New York: IEEE; 2011 Presented at: 2011 Winter Simulation Conference; 2011; Phoenix, AZ p. 2810-2821. [doi: 10.1109/WSC.2011.6147985]

17. Ratcliffe J, Simons A. How Can 3D Game Engines Create Photo-Realistic Interactive Architectural Visualizations? In: E-Learning and Games. Cham, Switzerland: Springer; 2017 Presented at: International Conference on Technologies for E-Learning and Digital Entertainment; 26-28 June 2017; Bournemouth, United Kingdom p. 1-11. [doi: 10.1007/978-3-319-65849-0_17]

18. Teyseyre A, Campo M. An Overview of 3D Software Visualization. IEEE Trans. Visual. Comput. Graphics 2009 Jan;15(1):87-105. [doi: 10.1109/tvcg.2008.86]

19. Zeckzer D, Müller R. Past, Present, and Future of 3D Software Visualization - A Systematic Literature Analysis. In: Proceedings of the 6th International Conference on Information Visualization Theory and Applications - Volume 1. Portugal: Science and Technology Publications, LDA; 2015 Mar Presented at: IVAPP 2015 - 6th International Conference on Information Visualization Theory and Applications; 2015; Berlin, Germany p. 63-74. [doi: 10.5220/0005325700630074]

20. Ploor DM. Video Game Design Foundations Software Design Guide. 2nd edition. Tinley Park, IL: Goodheart-Willcox Company; 2014:5-410.

21. Unity Game Engine. URL: https://unity3d.com/ [accessed 2020-04-18]

22. Adzharuddin N. Learning Management System (LMS) among University Students: Does It Work? IJEEEE 2013;3(3):248-252. [doi: 10.7763/ijeeee.2013.v3.233]

23. Wannapiroon P, Kaewrattanapat N, Premsmith J. Development of Cloud Learning Management Systems for Higher Education Institutions. In: 2019 Research, Invention, and Innovation Congress (RI2C). New York: IEEE; 2019 Presented at: 2019 Research, Invention, and Innovation Congress; 2019; Bangkok, Thailand p. 1-6. [doi: 10.1109/ri2c48728.2019.8999877]

24. Iqbal S. Learning Management Systems (LMS): Inside Matters. IMBR 2011 Oct 15;3(4):206-216. [doi: 10.22610/imbr.v3i4.935]

25. Bergeron B. Application of learning record stores and other forms of electronic competency records in modeling competency degradation. Stud Health Technol Inform 2014;196(23(8)):23-28. [Medline: 24732474]

26. Advanced Distributed Learning Initiative. URL: https://adlnet.gov [accessed 2019-01-07]

27. Schardosim SJL. Model for Recording Learning Experience Data from Remote Laboratories Using xAPI. In: Proceedings of the XIII Latin American Conference on Learning Technologies. New York: IEEE; 2018 Presented at: XIII Latin American Conference on Learning Technologies; 2018; Sao Paolo, Brazil p. 450-457. [doi: 10.1109/laclo.2018.00081]

28. MedCalc, version 12.7.7 [Statistical Software]. URL: https://www.medcalc.org/ [accessed 2020-08-21]

29. Creutzfeldt J, Hedman L, Felländer-Tsai L. Effects of pre-training using serious game technology on CPR performance-an exploratory quasi-experimental transfer study. Scand J Trauma Resusc Emerg Med 2012 Dec 06;20(1):79 [FREE Full text] [doi: 10.1186/1757-7241-20-79] [Medline: 23217084]

30. Semeraro F, Frisoli A, Ristagno G, Loconsole C, Marchetti L, Scapigliati A, et al. Relive: a serious game to learn how to save lives. Resuscitation 2014 Jul;85(7):e109-e110. [doi: 10.1016/j.resuscitation.2014.03.306] [Medline: 24709614]

XSL•FO

RenderX

31.   Boada I, Rodriguez-Benitez A, Garcia-Gonzalez JM, Olivet J, Carreras V, Sbert M. Using a serious game to complement
      CPR instruction in a nurse faculty. Comput Methods Programs Biomed 2015 Nov;122(2):282-291. [doi:
      10.1016/j.cmpb.2015.08.006] [Medline: 26319184]

Abbreviations

BLS: basic life support
LMS: learning management system
LRS: learning record store
SCORM: Sharable Content Object Reference Model
xAPI: experience API

XSL•FO
RenderX

Original Paper

Using Serious Games for Antismoking Health Campaigns: Experimental Study

Jihyun Kim[1], PhD; Hayeon Song[2], PhD; Kelly Merrill Jr[3], MA; Younbo Jung[4], PhD; Remi Junghuem Kwon[5], PhD

[1]University of Central Florida, Orlando, FL, United States

[2]Sungkyunkwan University, Seoul, Republic of Korea

[3]Ohio State University, Columbus, OH, United States

[4]Nanyang Technological University, Singapore, Singapore

[5]Korea University Technology and Education, Cheoan-si, Republic of Korea

Corresponding Author:
Hayeon Song, PhD
Sungkyunkwan University
25-2 Sungkyunkwan-ro, Jongno-gu
Seoul, 03063
Republic of Korea
Phone: 82 2 740 1870
Email: Hayeon.song@gmail.com

Related Article:

This is a corrected version. See correction statement: https://games.jmir.org/2021/1/e28180/

Abstract

Background: Serious games for health have been gaining in popularity among scholars and practitioners. However, there remain a few questions to be addressed.

Objective: This study tests the effects of a serious game and fear appeals on smoking-related outcomes. More specifically, this research aims to understand how serious games function as a more effective vehicle for a health campaign than a traditional medium, such as a print-based pamphlet. Further, while serious games utilize a variety of persuasive strategies in the game's content, it is not clear whether fear appeals, which are widely used persuasive-message strategies for health, can be an effective strategy in serious games. Thus, we are testing the effect of fear appeals in a serious game.

Methods: We created a computer game and a print brochure to educate participants about the risks of smoking. More specifically, a flash-based single-player game was developed in which players were asked to avoid cigarettes in the gameplay context. We also developed an online brochure based on existing smoking-related brochures at a university health center; antismoking messages on the computer game and in the brochure were comparable. Then, an experiment using a 2 (media type: game vs. print) x 2 (fearful image: fear vs. no-fear) between-subjects design was conducted. The study recruitment was announced to undergraduate students enrolled in a large, public Midwestern university in the United States. After a screening test, a total of 72 smokers, who reported smoking in the past 30 days, participated in the experiment.

Results: Overall, gameplay, when compared to print-based pamphlets, had greater impacts on attitudes toward smoking and the intention to quit smoking. Further, the game's persuasive effects were especially pronounced when messages contained fear appeals. When fearful images were presented, participants in the game condition reported significantly more negative attitudes toward social smoking than those in the print condition [$F(1,67)=7.28$; $P=.009$; $\eta_p^2=0.10$]. However, in the no-fear condition, there was no significant difference between the conditions [$F(1,67)=0.25$; $P=.620$]. Similarly, the intention to quit smoking [$F(1,67)=4.64$; $P=.035$; $\eta_p^2=0.07$] and susceptibility [$F(1,67)=6.92$; $P=.011$; $\eta_p^2=0.09$] were also significantly different between the conditions, but only when fear appeals were used.

Conclusions: This study extends fear appeal research by investigating the effects of different media types. It offers empirical evidence that a serious game can be an effective vehicle for fear appeals.

XSL•FO
RenderX

Introduction

Smoking and Smoking Cessation Interventions

Smoking is one of the most prevalent risky health behaviors. Currently, there are more than 35 million Americans that identify as smokers [1]. Although the number of smokers has consistently declined over the years, smoking has frequently been associated with negative health outcomes that often lead to death. Worldwide, smoking is responsible for more than 7 million deaths per year [2]. In the United States alone, more than 480,000 deaths a year are due to smoking, which includes 41,000 deaths due to second-hand smoking [3]. Researchers and medical professionals have also identified smoking as one of the most important risk factors for various health problems, such as different types of cancer [4], kidney failure [5], and cardiovascular disease [6]. In fact, more than 16 million Americans are living with a disease that was caused by smoking [3].

Smoking is particularly serious among young adults. Statistics indicate that nearly 9 out of 10 smokers try their first cigarette by the age of 18 years [3]. More than half of college smokers are reported as social smokers [7], and 1 in every 5 social smokers become a daily smoker during their 4 years at college [8]. In all, extant evidence suggests a strong need for investigating social smokers at the college level [9].

In an attempt to resolve prevalent smoking behaviors, much effort has focused on antismoking campaigns and interventions. For example, researchers have investigated the smoker identity to better understand how a smoker's self-identification is related to their attitude toward smoking [9]. Also, the effectiveness of diverse media types for smoking cessation interventions have been examined, such as text messaging [10] and web-based approaches [11]. However, regardless of these efforts, smoking issues remain. This calls for a need to consider a different type of medium for health intervention, such as serious games for health.

Serious games for health have been gaining popularity among scholars and practitioners. However, there remain a few questions to be addressed. First, there is a strong need to understand how serious games can be more persuasive compared to other media types, such as print pamphlets. Second, while serious games utilize a variety of persuasive strategies in the game content, it is not clear whether fear appeals, which are widely used persuasive-message strategies for health, can be an effective strategy in serious games. That is, collectively, little research has examined the effects of media types when fear appeals have been utilized in the game content. To fill in the gap, this study investigates the effectiveness of a serious game compared to a printed brochure in the context of fear appeals.

Fear Appeals

Fear appeals are described as "a persuasive message that attempts to arouse the emotion of fear by depicting a personally relevant and significant threat" [12]. That is, fear appeals are generally designed to scare people by describing negative outcomes that could occur when a recommended health behavior is not followed [12-14]. In acknowledgment of their effectiveness, fear appeals have been used as a core concept in a handful of health-related theories, such as the Protection Motivation Theory (PMT) [15] and the parallel process model [16]. Of the various fear-based theories, the extended parallel process model (EPPM) uses a more extensive approach to the understanding of fear [12] by combining aspects of PMT and the parallel process model. EPPM was developed to address the varying outcomes of these previous approaches regarding fear appeal messages. For example, research [12] emphasized "the role of the emotion fear in individuals' responses to fear appeals" in EPPM, which was not fully addressed in other theories.

Fear appeals have specifically been used to combat health issues and produce positive outcomes; fear-based messages have been incorporated in diverse health campaign materials to persuade people to change their behaviors and attitudes toward making healthy choices [17]. Song et al [18] found that seeing a visualized fearful future, which was presented via a deteriorated image of one's self due to smoking, increased not only smokers' negative attitudes toward social smoking but also their intention to stop smoking when compared to other smokers not exposed to the fearful future. Findings from a meta-analysis also noted the effectiveness of fear appeals [14].

Although fear has been studied for decades, questions remain. In particular, it is not clear whether a fear-based strategy would also work effectively in gaming contexts. The effects of fear appeals were examined in various media platforms (eg, text, graphic/images, video); however, the role of fear in interactive games (eg, serious games) compared to other traditional media platforms (eg, print-based brochure) has not been fully addressed. In this regard, there is a need to understand how a fear-based strategy would work in games, compared to other traditional media platforms.

Serious Games

Serious games are computer or video games designed for the primary purpose of educating users beyond entertainment purposes (eg, educational games, exercise video games, language learning games) [19-22]. Studies have demonstrated that serious games are effective in sharing information as well as changing attitudes and behaviors in a variety of health domains, such as safe sex [23], healthy eating [24-26], cancer [27,28], smoking [29,30], exercise [31-33], asthma management [34,35], and over-the-counter medication [36].

Several health education strategies are incorporated into serious games, including exposure therapy, behavioral rehearsal, and role-playing. The exposure therapy strategy provides repeated cues that trigger anxiety, often employed to diminish the anxiety or urge as a treatment for anxiety disorders such as acrophobia and arachnophobia [37-40]. The behavioral rehearsal strategy involves individuals practicing healthy behaviors, such as utilizing coping strategies to avoid relapse from smoking or

choosing healthy foods repeatedly in a game to encourage the participant to engage in similar behaviors in a real-life setting [26]. Lastly, the role-playing strategy is used to help people understand various social settings from other perspectives [41].

Features and Underlying Mechanisms of Serious Games

Games, compared to traditional media, have several unique built-in features. In particular, games can provide powerful learning experiences to change real-life behaviors with structural features and mechanisms that enable individuals to have their own experiences of the daunting future. In fact, research has identified various benefits of games for learning [42-47]. Specifically, Gee [43] highlighted certain features of games that are beneficial to learning. These include co-design and customization, where players have an interactive role with the content they engage with; manipulation, which allows players to engage in actions at a distance; and identity, where players have the ability to take on a new identity in games. These features allow the player to have new, meaningful experiences that have the potential to alter the player's thought processes. The following describe some unique features and underlying mechanisms of serious games in more detail: (1) the combination of enactive and vicarious learning; (2) persuasion through entertainment education; and (3) interactively tailored content.

Combination of Enactive and Vicarious Learning

Social Cognitive Theory [48] explains that there are 2 types of human learning processes: enactive and vicarious learning. Enactive learning occurs when individuals directly experience the consequences of their behaviors. Vicarious learning occurs when individuals observe other people's actions and thereby experience the consequences of their behaviors indirectly. Research [48,49] explains that enactive learning is more powerful than vicarious learning; however, enactive learning is often limited due to physical limitations (eg, time, space). Thus, vicarious learning is inevitable to expand humans' experiences beyond everyday life routines. Someone else's rare experiences or stories in the media (eg, movies, television) can provide a much wider range of occurrences that cannot be directly experienced.

Particularly, smoking-related fear appeals have been primarily associated with vicarious learning [50] because direct experiences of negative consequences could cause detrimental effects on the smoker's real body. Thus, efforts have been focused on sharing someone else's negative experiences of smoking via various methods such as images, messages, or testimonials. Research [49] argues that vicarious learning plays an important role by expanding knowledge and skills without necessarily having to participate directly in an action.

Following the above-mentioned argument, it is important to note that serious games can offer the benefits of both enactive and vicarious learning. First, the center of action resides in the game player rather than the observations of another individual's actions. During gameplay, a player uses an individualized avatar (ie, game characters), and the player's decisions and actions in the virtual environment have consequences for their avatar. In other words, players engage in enactive learning experiences

through the body of the avatar. This process, conceptualized as "mediated enactive experience" [26], is a unique feature of gameplay compared to other traditional media where individuals passively observe someone else's experiences. Additionally, Yee and Bailenson [51] found that people behave differently depending on the appearance of the avatar they control, which is understood as the Proteus Effect [51,52]. The Proteus Effect not only shows that the appearance of an avatar affects online behavior but also offline behavior [52]; an individual's behavior in the real world can be altered based on the appearance of an avatar they control in a digital world. Thus, in-game actions that alter an avatar's appearance not only influence the avatar but the individual controlling it as well.

Second, serious games also offer a breadth of vicarious learning. Unlike experiences in the physical environment, game experiences are not limited in terms of space, time, and the number of potential counterparts with whom players can interact. In gameplay, players can vicariously engage in actions and behaviors that may not be easily manifested in the physical world. Accordingly, a serious game is a unique vehicle; it provides a combination of both enactive and vicarious learning, can increase effective learning outcomes from the gameplay, and has the potential to be an effective persuasive medium.

Persuasion Through Entertainment Education

Fear appeals can be effective persuasive tools; however, they can also fail to accomplish what they were intended to do. One of the main reasons fear appeals can fail is that the experience of fear sometimes leads an individual to avoid or ignore a message that generates a negative emotion rather than changing risky behaviors [12]. Research has shown evidence that Entertainment-Education (E-E) can be successful by subverting topic avoidance caused by fear [53].

E-E is a strategy that utilizes popular entertainment media (eg, games, television, radio) to educate audiences [54,55]. The fun factor embedded in E-E materials generates intrinsic motivation and facilitates audience involvement that can result in persuasive message reception [56]. In a longitudinal qualitative study [57] focused on using a serious game to help with smoking cessation, participants indicated that they found the serious game to be fun; and these participants also demonstrated greater levels of motivation than the participants that were not exposed to a serious game.

For the same reason, people watch horror movies despite the negative emotions (eg, horror, sadness) they may experience, as the fun factor in the narratives holds an audience's attention through the duration of the message [53,58]. For example, a suspense study [58] showed individuals were willing to experience negative emotions and intensive arousal as a payoff to know what would happen next in the narratives. Previous work also explains that narrative forms of communication may sustain attention among low-motivation audiences and generate inadvertent persuasive effects [59]. Further, E-E has been found to be persuasive by reducing counterarguments [53]. According to the Extended Elaboration Likelihood Model (E-ELM), viewing dramatic elements in contexts (eg, gory and fearful images) leads to less critical perceptions of content [60-62]. However, viewers are less likely to develop counterarguments

as they encounter E-E messages and are therefore more likely to be persuaded by the messages.

Interactively Tailored Content

Messages can be more persuasive when they are tailored to the specific individual or group of people exposed to them. Tailoring is defined as "any combination of information or change strategies intended to reach one specific person based on characteristics unique to that person, related to the outcome of interest, and derived from an individual assessment" [63]. Thus, an advantage of tailored messages is their ability to depend on the individual differences among people, rather than assuming the effects are homogenous across a group of people [64].

Compared to nontailored messages, tailored messages are more likely to be read, understood, recalled, and perceived as trustworthy; thereby, they have a greater potential for persuasion and behavioral change [65-69]. Particularly, tailored messages have been effective in fear appeal strategies by motivating individuals to think that the risk condition may happen to them (ie, increasing perceived susceptibility) [70-72]. When tailored messages are provided, individuals tend to perceive that the messages are better fit for their condition [73,74], resulting in an increased level of perceived susceptibility. Important to note, however, is a meta-analysis using tailored health-behavior change interventions [75]. Results indicated that message tailoring has positive effects on most health-related attitudes except perceived susceptibility. While more research is needed to understand the results of the meta-analysis, a direction for future research would be testing the impact of tailored messages in varying degrees on perceived susceptibility.

As tailoring/customization is afforded in many digital and video games [76], serious games, too, can provide tailored health messages as well as personalized, interactive feedback associated with each action. For example, the game narratives develop based on a player's decisions and behaviors in a highly interactive fashion. As Rafaeli [77] indicated, interactivity, considered to be a natural characteristic of face-to-face conversation (eg, communicating with a personal counselor), is now commonly observed in the computer-mediated environment. Games have considerable opportunities for message tailoring and providing interactive feedback at any time with almost no additional cost, unlike traditional, print-based formats that may have limitations both in quantity and quality of message tailoring.

In fact, empirical research has found that people exposed to a more interactive message were more likely to change their attitudes [30]. For instance, participants exposed to serious games are able to abstain from smoking longer than participants that were not exposed to serious games [78,79]. This is likely due to gamified application, which helps to distract participants from smoking and provided helpful smoking cessation advice

[80]. Raiff et al [81] also found that participants are more likely to use a serious game to aid the cessation of smoking as compared to other cessation tactics, such as using a nicotine patch, a drug, or attending a support group.

This study compares the potential of a serious game to a traditional, print-based format regarding the use of fear appeals as a persuasive vehicle. More specifically, the current study hypothesizes that antismoking messages featured in a serious game will be more likely to positively influence smoking-related outcomes than a print-based brochure. Additionally, the study tests the role of fear appeals in serious games to understand whether media types would interplay with fear appeal strategies. To better understand the key aspects of smoking-related outcomes, the study specifically focuses on the following: attitudes toward smoking, the intention to quit smoking, and perceived susceptibility.

Based on the aforementioned literature, the following 3-part hypothesis is proposed: Individuals playing an antismoking game report stronger levels of (1) negative attitudes toward social smoking, (2) intention to quit smoking, and (3) perceived susceptibility to smoking-induced risks than individuals reading an antismoking print-based pamphlet. The following research question is also proposed: What role does a fear appeal strategy play in a serious game compared to a print brochure for an antismoking health campaign?

Methods

Participants

To identify eligible participants, a screening test was conducted online among undergraduate students enrolled in a large, public Midwestern university in the United States. Only individuals who reported smoking in the past 30 days were contacted. Thus, a total of 72 smokers were included in this study. Of the 72 participants, the sample comprised of 41 men (57%) and 31 women (43%), with a mean age of 21.40 (SD 5.14) years. The median smoking frequency was 2 to 3 times a week.

We conducted an experiment to test the hypothesis and research question. In particular, a 2 (media type: game vs. print) x 2 (fearful image: present vs. not present) between-subjects design was employed. Participants were assigned into 1 of the 4 experimental conditions: the game-fear condition (n=16), the game–no fear condition (n=16), the print-fear condition (n=19), and the print-no-fear condition (n=21). To ensure that there were no initial differences among the experimental groups, several analyses were conducted regarding sex, age, and smoking frequency. Results found that there were no statistically significant differences between groups. Thus, group equivalence was ensured. Demographic information for each experimental group is presented in Table 1.

Table 1. Description of the participants (n=72) in each experimental group.

Participant characteristics	Game-fear group (n=16)	Game–no fear group (n=16)	Print-fear group (n=19)	Print–no fear group (n=21)
Gender[a], n				
Male	11	9	9	12
Female	5	7	10	9
Age[b] in years, mean (SD)	19.81 (6.49)	21.88 (2.31)	22.58 (4.49)	21.19 (6.05)
Smoking frequency[c], median times per week	1	2-3	2-3	2-3

[a]Group difference for sex: $\chi^2(3)=1.624$; $P=.654$.

[b]Group difference for age: $F(3)=0.894$; $P=.449$.

[c]Group difference for smoking frequency: $\chi^2(18)=15.014$; $P=.661$.

Material Development

A series of steps were taken to develop materials for the experiment. First, a print brochure was created (Figure 1). To develop health messages for the brochure, all the antismoking related brochures available at the health center on the university campus were collected. Factoids, or information frequently appearing in the brochures, were identified and reformatted into a question-and-answer format containing 8 questions and answers. Then, a smoker's testimonial message, including feelings of physical weakness, regrets about not quitting smoking earlier, and complaints about looking old compared to nonsmokers of similar age, was created. The testimonial content was presented alongside a photo of the smoker in the brochure.

Fear conditions were manipulated based on the smoker's face. In the no-fear condition, only a young and healthy-looking face of a smoker was presented. However, in the fear condition, both a young, healthy-looking face of a smoker and an old, wrinkled face of the same smoker (which was caused by smoking) were presented.

In the fear condition for the print brochure, one page of the pamphlet showed a young and healthy-looking face with the subtitle, "Go to the next page to see how smoking can change her." On the following page, a discolored and wrinkled face was featured with a message stating, "Smoking can make you look 20 years older." Further explanation was provided: "Besides having the lungs of a senior, by the time a smoker turns 40, they will have approximately as many wrinkles as nonsmokers in their 60s." In the no-fear condition, only a young and healthy-looking face of a smoker was presented, and there was no mention of what smoking can change in a person's face. Otherwise, all the textual and graphical messages remained the same as the fear condition.

Figure 1. Antismoking intervention brochure. (The face of the avatar was not blurred in the original brochure; it has been blurred for this publication only).
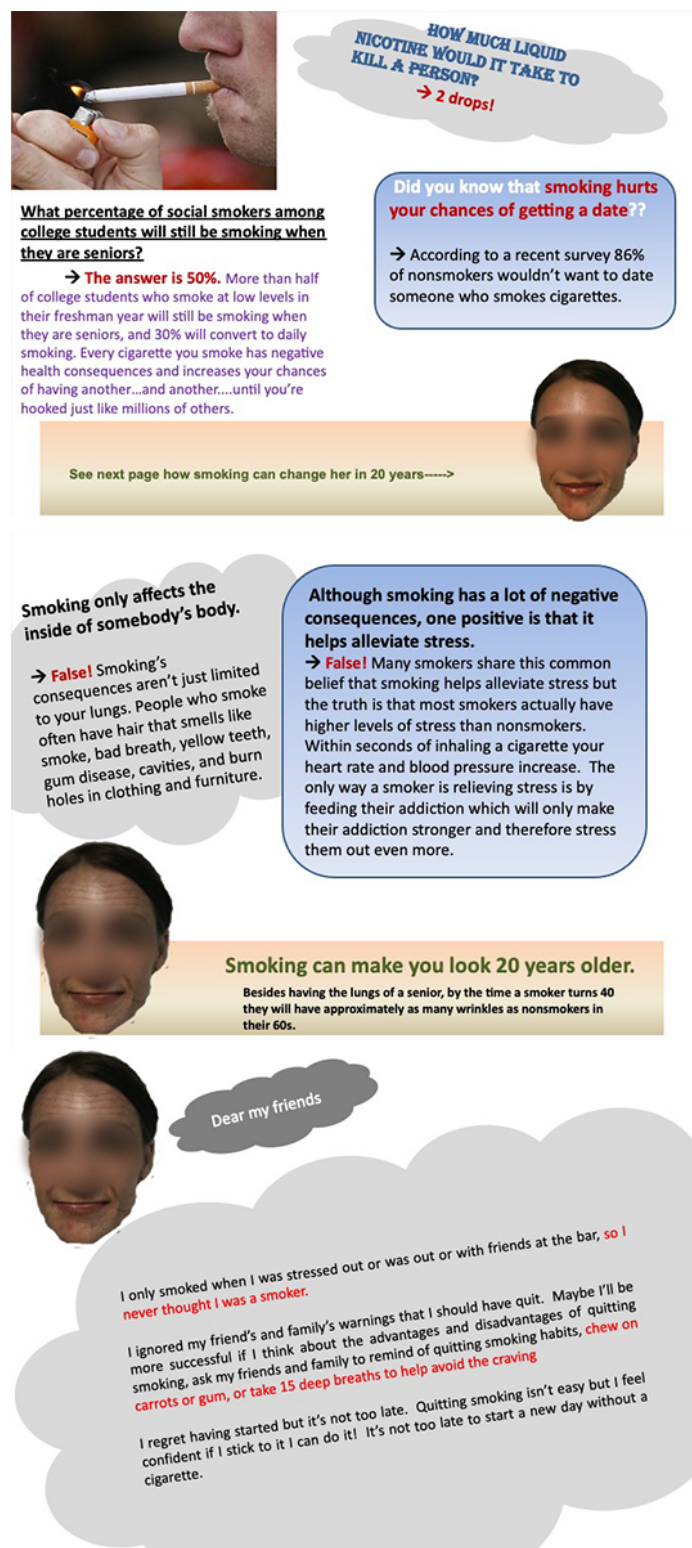
For the game condition (Figure 2), a computer game was developed for the purpose of educating smoking risks. The game is a flash-based single-player game where users are asked to avoid smoking cigarettes when they are stressed out because of an upcoming exam (Level 1) and when they are hanging out at a bar (Level 2). This game includes all 3 structural features of digital games that were discussed earlier (ie, combination of enactive and vicarious learning, entertainment education, and interactive tailoring). Players have full control of the avatar, enabling mediated enactive experiences. All of the actions (eg, avoiding smoking) are based on the player's own decisions. Players are in the center of the narratives and experience the negative consequences of the smokers (eg, not getting a second date, physical weakness, financial consequences) as their own experience rather than as someone else's (ie, combination of mediated enactive and vicarious experience). The game ends

with a monograph where the avatar regrets not quitting smoking while he or she was young. At the end of each level of the game, there is an educational quiz session, which was designed to be interactive. Depending on which answer the individual picks, true/false results are given, followed by a tailored explanation to give more detailed information about smoking. The game is designed to be easy to play—all of the participants in the game condition self-reported playing the game without difficulty.

Figure 2. Antismoking intervention computer game. (The face of the avatar was not blurred in the original game; it has been blurred for this publication only).
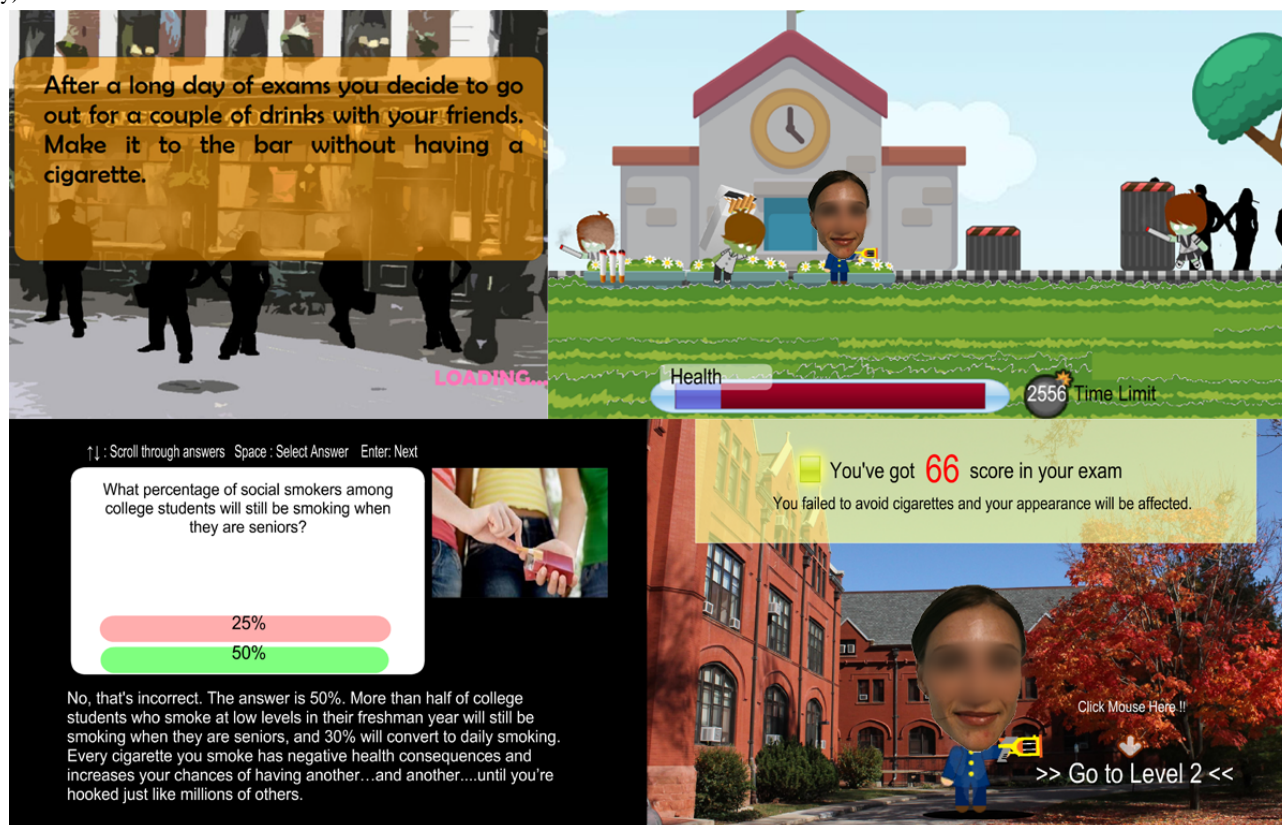

In the fear condition for the game, participants played with an avatar that had a young and healthy-looking face in Level 1. Then, they were told their avatar's face would look aged in Level 2 due to smoking experienced in Level 1. In the no-fear condition, a young and healthy-looking face was used in both Levels 1 and 2. In order to avoid any potential confounding effects, the gender of the smoker's face was matched with that of participants in both conditions. Also, the smoking-related content portrayed in the game was consistent with the print pamphlet.

Procedure

Upon the university's Institutional Review Board's approval, a recruitment message was distributed via the university's email system. Interested participants contacted the researcher via an online survey system. Prior to the experiment, researchers conducted a pretest to assess baseline smoking-related attitudes and perceptions. Approximately 2 weeks after the completion of the pretest, half of the participants, who were randomly selected to a game condition, were invited to a physical lab where they played the game. The other half of the participants were invited to an online lab where they viewed the pamphlet. Upon completing the main task of the study, participants were instructed to complete a posttest.

Measures

Pretest

Negative attitudes toward social smoking (α=.77) were measured by 6 items (eg, "Social smoking can cause health issues"). Intention to quit smoking (α=.89) was assessed using 8 items (eg, "I am willing to try quitting smoking"). All of the responses were obtained on a 10-point Likert-type scale (1=*Strongly Disagree*; 10=*Strongly Agree*).

Posttest

Negative attitudes toward social smoking (α=.77) and intention to quit smoking (α=.89) were measured in the posttest again, and the same items were used from the pretest. Susceptibility (α=.89) was measured by 6 items focusing on the negative consequences of smoking described in the antismoking messages (eg, "I think I will look old if I keep smoking" and "I know that I will have bad breath if I keep smoking"). Due to the specificity of the measure related to the messages in both the game and print conditions, considerable test-sensitization issues were expected. Therefore, susceptibility was asked only once in the posttest. All of the measures were developed for the study, and the responses were obtained on a 10-point Likert-type scale (1=*Strongly Disagree*; 10=*Strongly Agree*).

Results

Controlling for the sex of the participants, a series of analyses of covariance (ANCOVAs) were conducted using SPSS software (version 25; IBM Corp) to test the proposed hypothesis and the research question. Because attitudes towards social smoking and intention to quit smoking were measured at both pretests and posttests, they were tested based on the score difference between pretests and posttests (score=posttest – pretest). Therefore, positive scores indicate that posttest scores are higher than the pretest scores. Since susceptibility was measured in the posttest only, the final index was based on the posttest score.

Regarding the first part of the hypothesis (hypothesis 1), individuals in the game condition (mean 0.93, SD 1.60),

compared to those in the print condition (mean 0.09, SD 1.45), reported significantly more negative attitudes toward social smoking [$F(1,67)=5.25$; $P=.03$; $\eta_p^2=0.07$]. As for the second part of the hypothesis (hypothesis 2), there was no significant difference in the intention to quit smoking between the game condition (mean 0.49, SD 1.17) and the print condition (mean -0.22, SD 1.99) [$F(1,67)=2.83$; $P=.10$; $\eta_p^2=0.04$]. With regard to the third part of the hypothesis (hypothesis 3), individuals in the game condition (mean 8.48; SD 1.22) reported a greater level of susceptibility [$F(1,67)=5.81$; $P=.02$; $\eta_p^2=.08$] than those in the print condition (mean 7.65, SD 1.91). In sum, hypothesis 1 (attitudes) and hypothesis 3 (susceptibility) were supported, but hypothesis 2 (intention) was not. Table 2 illustrates the main effects of media format and fear.

Table 2. Main effects of media format and fear.

Condition	Condition, mean (SD)				Univariate F	P value	η_p^2
	Game	Print	Fear	No fear			
Media format							
Attitudes	0.93 (1.60)	0.09 (1.45)	N/A[a]	N/A	5.25	.03	.07
Intention	0.49 (1.17)	-0.22 (1.99)	N/A	N/A	2.83	.10	.04
Susceptibility	8.48 (1.22)	7.65 (1.91)	N/A	N/A	5.81	.02	.08
Fear							
Attitudes	N/A	N/A	0.60 (1.38)	0.33 (1.72)	0.78	.38	.01
Intention	N/A	N/A	0.15 (1.91)	0.05 (1.50)	0.13	.72	.002
Susceptibility	N/A	N/A	7.61 (2.01)	8.41 (1.21)	4.16	.045	.06

[a]N/A: not applicable.

To answer the research question that examined the interplay between media types and fear appeals, interaction effects and simple main effects were assessed (Table 3).

None of the interaction effects were significant [attitudes: $F(1,67)=2.52$, $P=.12$; intention: $F(1,67)=2.00$, $P=.16$; susceptibility: $F(1,67)=1.89$, $P=.17$]. However, the differences between the pretests and posttests appeared to be greater in the game condition compared to the print condition (Table 3).

Further analyses were conducted to examine the simple main effects of fear and media types. As 2 different comparisons were conducted in each simple test, the decision of statistical significance was made based on the adjusted P value .025 (.05/2) to reduce any potential error rate. First, when fearful images were presented, participants in the game condition reported significantly more negative attitudes toward social smoking than those in the print condition [$F(1,67)=7.28$; $P=.009$; $\eta_p^2=0.10$]. However, in the no-fear condition, there was no significant difference between the game and the print condition

[$F(1,67)=0.25$; $P=.62$]. Further, there was no significant difference between the fear and no-fear condition in the game condition [$F(1,67)=2.76$; $P=.10$] and in the print condition [$F(1,67)=0.29$; $P=.60$]. In all, the result indicated that the effect of fear on negative attitudes toward social smoking was stronger in the game condition than the print condition.

Next, simple main effects on the intention to quit smoking were tested. In the fear condition, participants in the game condition reported greater intention at a P value of .035 [$F(1,67)=4.64$; $P=.035$; $\eta_p^2=0.07$] compared to those in the print condition. However, in the no-fear condition, there was no significant difference [$F(1,67)=0.04$; $P=.85$]. Further, there were no significant differences between the fear and no-fear condition in the game condition [$F(1,67)=1.42$; $P=.24$] and in the print condition [$F(1,67)=0.63$; $P=.43$]. The result found that the effect of fear on the intention to quit smoking was stronger in the game condition than the print condition, although the P value was at .035.

Table 3. Simple effects of media format and fear.

Condition	Condition, mean (SD)				Univariate F	P value	η_p^2
	Game	Print	Fear	No fear			
Fear							
Attitudes	1.40 (1.02)	-0.06 (1.32)	N/A[a]	N/A	7.28	.009	.10
Intention	0.87 (1.23)	-0.46 (2.19)	N/A	N/A	4.64	.035	.07
Susceptibility	8.33 (1.52)	7.00 (2.20)	N/A	N/A	6.93	.011	.09
No fear							
Attitudes	0.47 (1.94)	0.22 (1.58)	N/A	N/A	0.25	.62	.004
Intention	0.11 (1.50)	0.01 (1.81)	N/A	N/A	0.04	.85	.001
Susceptibility	8.64 (0.84)	8.24 (1.42)	N/A	N/A	0.56	.46	.01
Game							
Attitudes	N/A	N/A	1.39 (1.02)	0.47 (1.94)	2.76	.10	.04
Intention	N/A	N/A	0.87 (1.23)	0.11 (1.01)	1.42	.24	.02
Susceptibility	N/A	N/A	8.33 (1.52)	8.64 (0.84)	0.19	.66	.003
Print							
Attitudes	N/A	N/A	-0.06 (1.32)	0.22 (1.58)	0.29	.60	.004
Intention	N/A	N/A	-0.46 (2.19)	0.01 (1.81)	0.63	.43	.01
Susceptibility	N/A	N/A	7.00 (2.20)	8.24 (1.42)	6.54	.013	.09

[a]N/A: not applicable.

Regarding susceptibility, when fear was presented, participants in the game condition reported greater susceptibility than those in the print condition [$F_{(1,67)}=6.92$; $P=.011$; $\eta_p^2= 0.09$]. However, when fear was not presented, there was no significant difference between the 2 media types [$F_{(1,67)}=0.56$; $P=.46$]. Further, in the game condition, there was no significant difference between the fear and no-fear condition [$F_{(1,67)}=0.19$; $P=.66$]. Finally, in the print condition, people in the no-fear condition reported greater susceptibility than those in the fear condition [$F_{(1,67)}=6.54$; $P=.013$; $\eta_p^2=0.09$]. The results indicated that the effect of fear on susceptibility was stronger in the game condition than the print condition. Further, in the print condition, not using fear was more effective than using fear in increasing susceptibility (Table 3).

Discussion

Principal Findings

This study compared the effectiveness of game- and print-based fear appeal messages in the context of social smoking, such as attitudes toward smoking, intention for smoking cessation, and susceptibility. Overall, the results indicate that games can be a more persuasive vehicle than print-based pamphlets when fear is incorporated in the message. Specifically, the study found that when smokers play the game, they experience more negative attitudes toward social smoking and greater susceptibility than those who read the pamphlet. Further, the result of the simple main effect analyses suggest that the game's persuasive effects are more pronounced when the media content includes fear messages; when a fearful image is presented in the form of an altered, aged face due to smoking, playing the game is significantly more likely to induce persuasive outcomes compared to reading the print pamphlet. However, such differences between the 2 media types are not observed when the fearful image is not presented.

Contributions and Implications

The present study extends fear appeal research by investigating the persuasive role of games and offers empirical evidence that a serious game can be an effective vehicle for fear appeal strategies. Games can provide unique opportunities that other traditional media cannot offer, and the use of games has the potential to further expand the fear appeal literature. For example, games can be an effective medium to identify and design an optimal situation to make fear appeal strategies work effectively, as game features make it possible to manipulate the characteristics of the avatar or narratives in order to enhance the effects of fear appeals. In addition, games can be tailored to each participant at reasonable costs. Serious games implemented in settings such as waiting rooms at hospitals and health care centers, for instance, may provide more successful outcomes than placing traditional pamphlets on magazine racks.

Another contribution of this study is the findings related to the effect of media types on susceptibility. Previous literature shows that promoting susceptibility has often been challenging because susceptibility is vulnerable to the innate human tendency to possess an optimistic bias [82]. Individuals are likely to prospect that their own future will be brighter compared to average people, which can interfere with a person's accurate assessment of health risks [83]. In a similar vein, a meta-analysis [75] revealed that tailored messages, which are widely considered to be an effective strategy in health campaigns, resulted in

XSL•FO
RenderX

positive changes for all other health-related attitudes except susceptibility. The meta-analysis study found that susceptibility demonstrated the opposite pattern; tailored messages reduced perceptions of susceptibility. A plausible explanation is that the attempt to tailor messages in traditional formats of media is limited in its ability to trigger enactive learning experiences, which sometimes results in the boomerang effect of reduced susceptibility. However, the findings of our study imply that serious games may enhance the perception of susceptibility more than print-based brochures. In fact, this finding is supported by research that interactive media enhances perceived susceptibility [71]. As such, this study argues that the interplay among susceptibility, media types, and fear appeals is important to note and calls for more research.

In addition, this study further strengthens the extant body of research. While a handful of studies have already investigated serious games and smoking, the current investigation is unique from the others in various ways. Specifically, this study examines additional key variables, as compared to the previous studies. In particular, one of the foci of this investigation is on susceptibility, which has received relatively little attention in previous game studies, except in the Song et al [18] study. Also, the present investigation has more potential to generalize the findings to a broader population. While most of the studies similarly indicated that participants had to be at least 18 years old and active smokers, some focused on specific populations, such as pregnant women [79], previously diagnosed cancer patients [78], and individuals identified as having a mental illness [84]. Some studies employed even more requirements. For example, participants had to either own a smartphone [57,79], score above a specific criterion for motivation to stop smoking [79], or had to be currently taking psychiatric medicine, as well as have an assigned psychiatric case manager and psychiatric provider [84]. While these studies provide specific information for particular groups of smokers, the implications of the study interventions might be somewhat limited to a targeted group. In this regard, this study has the potential to generalize results more broadly, as participants were not limited to particular conditions. Moreover, this study effectively compares differences in key variables between a printed pamphlet and a serious game. While some studies (such as the one conducted by Vilardaga et al [84]) tested a paper prototype of their serious game to aid in the development of their actual serious game, the comparison of a serious game and some form of a paper pamphlet is largely missing from previous studies.

Furthermore, this study offers implications for practice. Although an increasing number of studies have reported the effectiveness of games, relatively few have compared the effects of serious games for health education compared to the effects of traditional, printed brochures in the context of fear messages. In this regard, this study demonstrates empirical evidence that health practitioners, especially those who are planning to use fear appeal strategies, should pay close attention to interactive media such as serious games. The study's findings imply that the threat induced by a comparable description of the dire consequences of current health behaviors can be used more effectively to change problematic behaviors when a game, compared to printed brochures, is utilized. Games can provide

powerful experiences to change real-life behaviors with their structural features [85], which can also enable individuals to have their own experience of the daunting future that has not yet materialized in a highly tailored and fun way, paradoxically.

Additionally, this study suggests the adoption of serious games, especially when the intervention is targeting individuals with a low-risk perception. For example, social smokers do not tend to believe that they will suffer from the negative consequences of smoking and even fail to identify themselves as smokers [86]. Thus, interventions targeting regular or habitual smokers are easily neglected by social or occasional smokers. In this regard, games tend to be more approachable to them, and they can be more persuasive.

However, it is important to note that this study does not always suggest using serious games. As part of this study's findings indicate, there was no significant difference between the 2 media types in the no-fear condition. This result implies that different media types are likely to provide maximum effectiveness when their unique strengths are considered in developing campaigns. For example, print brochures may be useful for providing simple information about health behaviors, while serious games may be helpful for increasing perceived susceptibility in fear appeal messages. It is also important to acknowledge cost-related issues. To avoid offsetting monetary challenges, one could use print brochures to provide simple information about health behaviors for a relatively cheap cost. However, serious games, compared to printed brochures, may be more helpful for increasing perceived susceptibility in fear appeal messages. Then, a free app on a smartphone could make serious games more accessible to the general public. Future research needs to explore various ways to utilize diverse media types in order to maximize the effectiveness of health campaigns.

Limitations and Future Research Directions

As with any research, this study also has a few limitations that should be considered when interpreting the results. First, the study utilized fearful images to investigate the effect of fear appeals but did not consider other types of fear appeal strategies. Given that fear appeals are multifaceted stimuli [15], research has indicated a variety of factors that may influence the effects of fear appeal strategies such as narratives. In order to further expand the study's findings, future research should investigate various stimuli or the combination of multiple stimuli for fear appeals. In that way, the findings may help design more effective fear appeal strategies by providing detailed practical information.

Second, the study acknowledges that different ranges of exposure time to the study's stimulus may have influenced the results. Although the message presented in each medium (game and print) was comparable, it was not possible to control the time that each participant spent in their experimental condition. In particular, due to the nature of the medium, participants in the game condition may have spent a bit more time completing their tasks than those in the print condition. Even in the game condition, some participants may have finished the gameplay faster than others. In this regard, future researchers are encouraged to further investigate this issue.

Next, participants in this study were all college students. While previous research has already identified social smoking as being a prevalent behavior among college students [7-9], many others are engaging in this risky health behavior [78,79,84]. Thus, it is possible that the results of the study might be limited to this college student population. Future research should incorporate a more representative sample to ensure these results are more generalizable.

Moreover, the study focused on short-term effects only. Participants were asked their attitudes toward smoking only after interacting with either the game or the pamphlet. Thus, it is not clear whether these interactions would have a long-term effect on the participant's behavior. Future research should conduct a longitudinal study to assess how effective games and pamphlets are at inducing behavioral changes, such as intentions to quit smoking, over time.

Lastly, although this study did not investigate the role of narratives specifically, the study acknowledges that narratives in games play an important role, especially in persuasion [87]. With the coined term "procedural rhetoric," Bogost [87] explains that the users of persuasive games learn through the authorship of rules and processes by becoming the game character. Bogost [87] argues that learning achieved by experiencing the procedural rhetoric in persuasive games is quite different from learning-based simply on words/texts and visual images. Similarly, the concept of narrative persuasion has also been conceptualized and tested together with the feeling of transportation in persuasion and education contexts [88,89]. While these concepts may have played a role in this study's findings, they were not empirically tested in this investigation. Future researchers are encouraged to further expand this line of research.

Conclusion

Games have great potential to expand fear appeal research with their unique features. As with media convergence, games are gradually merged into other media such as social media and social television. Thus, it is important to further investigate ways to utilize games merged into other media to expand the fear appeal literature. In this regard, this study sheds light on the use of serious games in the context of utilizing fear appeals.

Conflicts of Interest

None declared.

References

1. Smoking of cigarettes in the U.S. 2019. Statista. 2019. URL: https://www.statista.com/statistics/275975/us-households-smoking-of-cigarettes/
2. WHO report on the global tobacco epidemic 2017. World Health Organization. 2017. URL: https://www.who.int/tobacco/global_report/2017/en/
3. 50 years of progress: a report of the surgeon general. U.S. Department of Health and Human Services. 2014. URL: https://www.hhs.gov/surgeongeneral/reports-and-publications/tobacco/consequences-smoking-factsheet/index.html
4. Cancer facts and figures 2018. American Cancer Society. 2018. URL: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2018/cancer-facts-and-figures-2018.pdf
5. Orth S. Smoking and the Kidney. Journal of the American Society of Nephrology 2002;13(6):1663-1672 [FREE Full text] [doi: 10.1097/01.ASN.0000018401.82863.FD]
6. Burns DM. Epidemiology of smoking-induced cardiovascular disease. Prog Cardiovasc Dis 2003;46(1):11-29. [doi: 10.1016/s0033-0620(03)00079-3] [Medline: 12920698]
7. Choi Y, Choi SM, Rifon N. "I smoke but I am not a smoker": phantom smokers and the discrepancy between self-identity and behavior. J Am Coll Health 2010;59(2):117-125. [doi: 10.1080/07448481.2010.483704]
8. Kenford SL, Wetter DW, Welsch SK, Smith SS, Fiore MC, Baker TB. Progression of college-age cigarette samplers: what influences outcome. Addict Behav 2005;30(2):285-294. [doi: 10.1016/j.addbeh.2004.05.017] [Medline: 15621399]
9. Song H, Kim J, Kim S. Smoker identity among social smokers: theory-based approaches for anti-smoking interventions. Journal of Substance Use 2013;19(5):346-350. [doi: 10.3109/14659891.2013.820799]
10. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS Text Message Interventions for Smoking Cessation: A Meta-Analysis. J Subst Abuse Treat 2015;56:1-10. [doi: 10.1016/j.jsat.2015.01.011] [Medline: 25720333]

11. Macy JT, Chassin L, Presson CC, Sherman JW. Changing implicit attitudes toward smoking: results from a web-based approach-avoidance practice intervention. J Behav Med 2015;38(1):143-152 [FREE Full text] [doi: 10.1007/s10865-014-9585-2] [Medline: 25059750]

12. Witte K. Putting the fear back into fear appeals: The extended parallel process model. Communication Monographs 1992 Dec;59(4):329-349. [doi: 10.1080/03637759209376276]

13. Rogers R. Cognitivepsychological processes in fear appealsattitude change: A revised theory of protection motivation. Inocial Psychophysiology: A sourcebook. Cacioppo J, Petty R. eds. , New York, NY: Guilford Press; 1983:A.

14. Witte K, Allen M. A meta-analysis of fear appeals: implications for effective public health campaigns. Health Educ Behav 2000;27(5):591-615. [doi: 10.1177/109019810002700506] [Medline: 11009129]

15. Rogers RW. A Protection Motivation Theory of Fear Appeals and Attitude Change1. J Psychol 1975;91(1):93-114. [doi: 10.1080/00223980.1975.9915803] [Medline: 28136248]

16. Leventhal H. Findings and theory in the study of fear communications. Adv Exp Soc Psychol 1970;5:119-186.

17. Roskos‑Ewoldsen DR, Yu JH, Rhodes N. Fear appeal messages affect accessibility of attitudes toward the threat and adaptive behaviors. Communication Monographs 2004;71(1):49-69. [doi: 10.1080/0363452042000228559]

18. Song H, Kim J, Kwon R, Jung Y. Anti-smoking educational game using avatars as visualized possible selves. Computers in Human Behavior 2013;29(5):2029-2036 [FREE Full text] [doi: 10.1016/j.chb.2013.04.008]

19. Charsky D. From Edutainment to Serious Games: A Change in the Use of Game Characteristics. Games and Culture 2010;5(2):177-198. [doi: 10.1177/1555412009354727]

20. Courbet D, Bernard F, Joule R, Halimi-Falkowicz S, Guéguen N. Small clicks, great effects: the immediate and delayed influence of websites containing serious games on behavior and attitude. International Journal of Advertising 2015;35(6):949-969. [doi: 10.1080/02650487.2015.1082226]

21. Meijer HA, Graafland M, Goslings JC, Schijven MP. Systematic Review on the Effects of Serious Games and Wearable Technology Used in Rehabilitation of Patients With Traumatic Bone and Soft Tissue Injuries. Arch Phys Med Rehabil 2018;99(9):1890-1899. [doi: 10.1016/j.apmr.2017.10.018] [Medline: 29138050]

22. Serious games. Sawyer B. 2006. URL: http://www.seriousgames.org./index.html

23. Ndulue C, Orji R. STD Pong: changing risky sexual behaviour in Africa through persuasive games. Paper presented at: Proceedings of the Second African Conference for Human Computer Interaction: Thriving Communities 2018; 2018 Presented at: The Second African Conference for Human Computer Interaction; December 3–7; Windhoek, Namibia. [doi: 10.1145/3283458.3283463]

24. Baranowski T, Baranowski J, Cullen KW, Marsh T, Islam N, Zakeri I, et al. Squire's Quest! Dietary outcome evaluation of a multimedia game. Am J Prev Med 2003;24(1):52-61. [doi: 10.1016/s0749-3797(02)00570-6] [Medline: 12554024]

25. Dominguez-Rodriguez A, Oliver E, Cebolla A. Serious Games to Teach Nutrition Education to Children Between 9 to 12 Years Old. Pickit! and Cookit!. eHealth 360°. ISBN 2017:143-147. [doi: 10.1007/978-3-319-49655-9_19]

26. Peng W. The mediational role of identification in the relationship between experience mode and self-efficacy: Enactive role-playing versus passive observation. Cyberpsychol Behav 2008;11(6):649-652. [doi: 10.1089/cpb.2007.0229] [Medline: 18954275]

27. Beale IL, Kato PM, Marin-Bowling VM, Guthrie N, Cole SW. Improvement in cancer-related knowledge following use of a psychoeducational video game for adolescents and young adults with cancer. J Adolesc Health 2007;41(3):263-270. [doi: 10.1016/j.jadohealth.2007.04.006] [Medline: 17707296]

28. De la Hera Conde-Pumpido T. The Persuasive Roles of Digital Games: The Case of Cancer Games. Media and Communication 2018;6(2):103-111. [doi: 10.17645/mac.v6i2.1336]

29. Song H, Kim J, Kwon RJ, Jung Y. Anti-smoking educational game using avatars as visualized possible selves. Computers in Human Behavior 2013;29(5):2029-2036. [doi: 10.1016/j.chb.2013.04.008]

30. Yoo S, Eastin M. Designing Health Games for Anti-Smoking Advertising Targeting College Students: The Impact of Message Types and Voice-Over. Int J Contents 2017;13(3):17-24. [doi: 10.5392/IJoC.2017.13.3.017]

31. Kim J, Timmerman CE. Effects of Supportive Feedback Messages on Exergame Experiences. Journal of Media Psychology 2018;30(1):29-40. [doi: 10.1027/1864-1105/a000175]

32. Song H, Kim J, Lee KM. Virtual vs. real body in exergames: Reducing social physique anxiety in exercise experiences. Computers in Human Behavior 2014;36:282-285. [doi: 10.1016/j.chb.2014.03.059]

33. Song H, Kim J, Tenzek KE, Lee KM. The effects of competition and competitiveness upon intrinsic motivation in exergames. Computers in Human Behavior 2013;29(4):1702-1708. [doi: 10.1016/j.chb.2013.01.042]

34. Drummond D, Monnier D, Tesnière A, Hadchouel A. A systematic review of serious games in asthma education. Pediatr Allergy Immunol 2017;28(3):257-265. [doi: 10.1111/pai.12690] [Medline: 27992659]

35. Homer C, Susskind O, Alpert H. An evaluation of an innovative multimedia educational software program for asthma management: report of a randomized, controlled trial. Pediatrics 2000;106(1):210-215 [FREE Full text]

36. Abraham O, Feathers A, Grieve L, Babichenko D. Developing and piloting a serious game to educate children about over‑the‑counter medication safety. J Pharm Health Serv Res 2019;10(2):235-241. [doi: 10.1111/jphs.12292]

37. Bordnick PS, Graap KM, Copp HL, Brooks J, Ferrer M. Virtual reality cue reactivity assessment in cigarette smokers. Cyberpsychol Behav 2005;8(5):487-492. [doi: 10.1089/cpb.2005.8.487] [Medline: 16232041]

38.   Bordnick PS, Traylor A, Copp HL, Graap KM, Carter B, Ferrer M, et al. Assessing reactivity to virtual reality alcohol based cues. Addict Behav 2008;33(6):743-756. [doi: 10.1016/j.addbeh.2007.12.010] [Medline: 18282663]

39.   Choi J, Park S, Lee J, Jung H, Lee H, Jin C, et al. The effect of repeated virtual nicotine cue exposure therapy on the psychophysiological responses: a preliminary study. Psychiatry Investig 2011;8(2):155-160 [FREE Full text] [doi: 10.4306/pi.2011.8.2.155] [Medline: 21852993]

40.   Miloff A, Lindner P, Hamilton W, Reuterskiöld L, Andersson G, Carlbring P. Single-session gamified virtual reality exposure therapy for spider phobia vs. traditional exposure therapy: study protocol for a randomized controlled non-inferiority trial. Trials 2016;17:60 [FREE Full text] [doi: 10.1186/s13063-016-1171-1] [Medline: 26833396]

41.   Melissourgos A, Paraskeva F. In a Q-Quest to Learn: A Collaborative Game That Supports Evaluation. 2017 Presented at: European Conference on Games Based Learning; Oct 5-6, 2017; Graz, Austria.

42.   Gee J. What video games have to teach us about learning and literacy. Comput. Entertain 2003;1(1):20 [FREE Full text] [doi: 10.1145/950566.950595]

43.   Gee J. Learning by Design: Good Video Games as Learning Machines. E-Learning and Digital Media 2016;2(1):5-16 [FREE Full text] [doi: 10.2304/elea.2005.2.1.5]

44.   Gee J. Are video games good for learning? Nordic Journal of Digital Literacy 2006;1(03):172-183.

45.   Freire M, Serrano-Laguna , Manero B, Martínez-Ortiz I, Moreno-Ger P, Fernández-Manjón B. Game learning analytics: learning analytics for serious games. In: Learning, design, and technology. Springer Nature Switzerland AG; 2016:1-29.

46.   Shaffer DW, Squire KR, Halverson R, Gee JP. Video Games and the Future of Learning. Phi Delta Kappan 2016;87(2):105-111. [doi: 10.1177/003172170508700205]

47.   Squire K. Video games and learning: Teaching and participatory culture in the digital age. New York, NY: Teachers College Press; 2011.

48.   Bandura A, Freeman W, Lightsey R. Self-efficacy: The exercise of control. New York, NY: Macmillan; 1997.

49.   Bandura A. Social Cognitive Theory of Mass Communication. Media Psychology 2001;3(3):265-299. [doi: 10.1207/S1532785XMEP0303_03]

50.   Malouff J, Thorsteinsson E, Schutte N, Rooke SE. Effects of Vicarious Punishment: A Meta-Analysis. The Journal of General Psychology 2009;136(3):271-286. [doi: 10.3200/genp.136.3.271-286]

51.   Yee N, Bailenson J. The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. Human Comm Res 2007;33(3):271-290 [FREE Full text] [doi: 10.1111/j.1468-2958.2007.00299.x]

52.   Yee N, Bailenson J, Ducheneaut N. The Proteus Effect. Communication Research 2009;36(2):285-312 [FREE Full text] [doi: 10.1177/0093650208330254]

53.   Moyer-Gusé E. Toward a theory of entertainment persuasionxplaining the persuasive effects of entertainment-education messages. Communication Theory 2008;18(3). [doi: 10.1111/j.1468-2885.2008.00328.x]

54.   Singhal A, Rogers E. The entertainment-education strategy in communication campaigns. In: Rice RE, Atkin CK, editors. Public Communication Campaigns. Thousand Oaks, CA: Sage; 2001:343-356.

55.   Behm-Morawitz E, Aubrey JS, Pennell H, Kim KB. Examining the Effects of MTV's 16 and Pregnant on Adolescent Girls' Sexual Health: The Implications of Character Affinity, Pregnancy Risk Factors, and Health Literacy on Message Effectiveness. Health Commun 2019;34(2):180-190. [doi: 10.1080/10410236.2017.1399506] [Medline: 29125323]

56.   Singhal A, Cody M, Rogers E, Sabido M. Entertainment-education and social change: History, research, and practice. Mahwah, NJ: Lawrence Erlbaum Inc; 2003.

57.   El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game On? Smoking Cessation Through the Gamification of mHealth: A Longitudinal Qualitative Study. JMIR Serious Games 2016;4(2):e18 [FREE Full text] [doi: 10.2196/games.5678] [Medline: 27777216]

58.   Zillmann D. The psychology of suspense in dramatic exposition. In: Vorderer P, Wulff H, Friedrichsen M, editors. Suspense: Conceptualizations, theoretical analyses, and empirical explorations. Mahwah, NJ: Erlbaum Inc; 1996:199-232.

59.   Green M. Narratives and cancer communication. Journal of Communication 2006;56:163-183. [doi: 10.1111/j.1460-2466.2006.00288.x]

60.   Slater M. Entertainment education and the persuasive impact of narratives. In: Green MC, Strange JJ, Brock TC, editors. Narrative impact: Social and cognitive foundations. Mahwah, NJ: Lawrence Erlbaum Inc; 2002:157-182.

61.   Slater M. Involvement as goal-directed strategic processing: Extending the elaboration likelihood model. In: Dillard JP, Pfau M, editors. The persuasion handbook: Developments in theory and practice. Thousand Oaks, CA: Sage; 2002:175-194.

62.   Slater MD, Rouner D. Entertainment?Education and Elaboration Likelihood: Understanding the Processing of Narrative Persuasion. Communication Theory 2002;12(2):173-191. [doi: 10.1111/j.1468-2885.2002.tb00265.x]

63.   Kreuter MW, Bull FC, Clark EM, Oswald DL. Understanding how people process health information: a comparison of tailored and nontailored weight-loss materials. Health Psychol 1999;18(5):487-494. [doi: 10.1037//0278-6133.18.5.487] [Medline: 10519465]

64.   Lucas T, Manning M, Hayman LW, Blessman J. Targeting and tailoring message-framing: the moderating effect of racial identity on receptivity to colorectal cancer screening among African-Americans. J Behav Med 2018;41(6):747-756. [doi: 10.1007/s10865-018-9933-8] [Medline: 29882172]

65.   Cheung K, Wijnen BFM, Hiligsmann M, Coyle K, Coyle D, Pokhrel S, et al. Is it cost-effective to provide internet-based
       interventions to complement the current provision of smoking cessation services in the Netherlands? An analysis based on
       the EQUIPTMOD. Addiction 2018;113:87-95 [FREE Full text] [doi: 10.1111/add.14069] [Medline: 29243351]

66.   Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring health messages: Customizing communication with computer
       technology. New York, NY: Routledge; 2013.

67.   Kreuter MW, Holt CL. How Do People Process Health Information? Applications in an Age of Individualized
       Communication. Curr Dir Psychol Sci 2016;10(6):206-209. [doi: 10.1111/1467-8721.00150]

68.   Rimal R, Adkins A. Using computers to narrowcast health messages: The role of audience segmentation, targeting,tailoring
       in health promotion. In: Thompson TL, Dorsey A, Miller K, Parrott R, editors. Handbook of Health Communication.
       Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2003:497-513.

69.   Smit E, Linn A. Julia Cv. Taking online computer-tailoring forward. Eur Health Psychol 2013;17(1):25-31.

70.   Campbell MK, DeVellis BM, Strecher VJ, Ammerman AS, DeVellis RF, Sandler RS. Improving dietary behavior: the
       effectiveness of tailored messages in primary care settings. Am J Public Health 1994;84(5):783-787. [doi:
       10.2105/ajph.84.5.783] [Medline: 8179049]

71.   Pot M, Van Keulen H, Paulussen T, Otten W, Van Steenbergen J, Ruiter R. Mothers' Perceptions of their Daughters'
       Susceptibility to HPV-related Risk Factors: An Experimental Pretest Comparing Narrative and Statistical Risk Information.
       Health Psychol Bull 2019;3(1):38. [doi: 10.5334/hpb.7]

72.   Rimal RN, Morrison D. A uniqueness to personal threat (UPT) hypothesis: how similarity affects perceptions of susceptibility
       and severity in risk assessment. Health Communication 2006;20(3):209-219. [doi: 10.1207/s15327027hc2003_1] [Medline:
       17137413]

73.   Kreuter MW, Oswald DL, Bull FC, Clark EM. Are tailored health education materials always more effective than non-tailored
       materials? Health Educ Res 2000;15(3):305-315. [doi: 10.1093/her/15.3.305] [Medline: 10977378]

74.   Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am
       J Health Behav 2003;27:227-232. [doi: 10.5993/ajhb.27.1.s3.6] [Medline: 14672383]

75.   Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change
       interventions. Psychol Bull 2007;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]

76.   Yee N. Befriending ogres and wood-elves: Relationship formation and the social architecture of Norrath. Game Studies
       2009;9(1).

77.   Rafaeli S. Interactivity: From new media to communication. In: Annual Review of Communication Research: Advancing
       Communication Science. Sage: Beverly Hills, CA; 1988:110-134.

78.   Krebs P, Burkhalter J, Fiske J, Snow H, Schofield E, Iocolano M, et al. The QuitIT Coping Skills Game for Promoting
       Tobacco Cessation Among Smokers Diagnosed With Cancer: Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth
       2019;7(1):e10071 [FREE Full text] [doi: 10.2196/10071] [Medline: 30632971]

79.   Marin-Gomez FX, Garcia-Moreno Marchán R, Mayos-Fernandez A, Flores-Mateo G, Granado-Font E, Barrera Uriarte
       ML, et al. Exploring Efficacy of a Serious Game (Tobbstop) for Smoking Cessation During Pregnancy: Randomized
       Controlled Trial. JMIR Serious Games 2019;7(1):e12835 [FREE Full text] [doi: 10.2196/12835] [Medline: 30916655]

80.   Edwards EA, Caton H, Lumsden J, Rivas C, Steed L, Pirunsarn Y, et al. Creating a Theoretically Grounded, Gamified
       Health App: Lessons From Developing the Cigbreak Smoking Cessation Mobile Phone Game. JMIR Serious Games
       2018;6(4):e10252 [FREE Full text] [doi: 10.2196/10252] [Medline: 30497994]

81.   Raiff BR, Fortugno N, Scherlis DR, Rapoza D. A Mobile Game to Support Smoking Cessation: Prototype Assessment.
       JMIR Serious Games 2018;6(2):e11 [FREE Full text] [doi: 10.2196/games.9599] [Medline: 29880466]

82.   Weinstein ND. Unrealistic optimism about future life events. Journal of Personality and Social Psychology
       1980;39(5):806-820. [doi: 10.1037/0022-3514.39.5.806]

83.   Radcliffe NM, Klein WMP. Dispositional, Unrealistic, and Comparative Optimism: Differential Relations with the Knowledge
       and Processing of Risk Information and Beliefs about Personal Risk. Pers Soc Psychol Bull 2016;28(6):836-846. [doi:
       10.1177/0146167202289012]

84.   Vilardaga R, Rizo J, Zeng E, Kientz JA, Ries R, Otis C, et al. User-Centered Design of Learn to Quit, a Smoking Cessation
       Smartphone App for People With Serious Mental Illness. JMIR Serious Games 2018;6(1):e2 [FREE Full text] [doi:
       10.2196/games.8881] [Medline: 29339346]

85.   Baranowski T, Buday R, Thompson D, Lyons EJ, Lu AS, Baranowski J. Developing Games for Health Behavior Change:
       Getting Started. Games Health J 2013;2(4):183-190 [FREE Full text] [doi: 10.1089/g4h.2013.0048] [Medline: 24443708]

86.   Berg CJ, Lust KA, Sanem JR, Kirch MA, Rudie M, Ehlinger E, et al. Smoker self-identification versus recent smoking
       among college students. Am J Prev Med 2009;36(4):333-336. [doi: 10.1016/j.amepre.2008.11.010] [Medline: 19201148]

87.   Bogost I. Persuasive games: The expressive power of videogames. Cambridge, MA: MIT Press; 2010.

88.   Green M, Brock T. In the mind's eye: Transportation-imagery model of narrative persuasion. In: Green M, Strange J, Brock
       T, editors. Narrative impact: Social and cognitive foundations. New York, NY: Psychology Press; 2002:315-341.

89.   Slater MD, Rouner D. Entertainment?Education and Elaboration Likelihood: Understanding the Processing of Narrative
       Persuasion. Communication Theory 2002;12(2):173-191. [doi: 10.1111/j.1468-2885.2002.tb00265.x]

Abbreviations

E-E: Entertainment-Education
EPPM: extended parallel process model
PMT: Protection Motivation Theory

XSL•FO
RenderX

Original Paper

Learning Impact of a Virtual Brain Electrical Activity Simulator Among Neurophysiology Students: Mixed-Methods Intervention Study

Marko Henrik Björn[1*], MSc; Jonne MM Laurila[2*], PhD; Werner Ravyse[3*], MSc; Jari Kukkonen[1*], PhD; Sanna Leivo[4*], MSc; Kati Mäkitalo[5*], Prof Dr; Tuula Keinonen[1*], Prof Dr

[1]School of Applied Educational Science and Teacher Education, Joensuu, University of Eastern Finland, Joensuu, Finland

[2]Institute of Biomedicine, Integrative Physiology and Pharmacology, University of Turku, Turku, Finland

[3]Turku University of Applied Sciences, Turku, Finland

[4]Department of Clinical Neurophysiology, Turku University Hospital, Turku, Finland

[5]Faculty of Education, University of Oulu, Oulu, Finland

[*]all authors contributed equally

Corresponding Author:
Marko Henrik Björn, MSc
School of Applied Educational Science and Teacher Education, Joensuu
University of Eastern Finland
Futura
Yliopistonkatu 7
Joensuu, PL 111
Finland
Phone: 358 0400587693
Email: marko.bjorn@turkuamk.fi

Abstract

Background: Virtual simulation is the re-creation of reality depicted on a computer screen. It offers the possibility to exercise motor and psychomotor skills. In biomedical and medical education, there is an attempt to find new ways to support students' learning in neurophysiology. Traditionally, recording electroencephalography (EEG) has been learned through practical hands-on exercises. To date, virtual simulations of EEG measurements have not been used.

Objective: This study aimed to examine the development of students' theoretical knowledge and practical skills in the EEG measurement when using a virtual EEG simulator in biomedical laboratory science in the context of a neurophysiology course.

Methods: A computer-based EEG simulator was created. The simulator allowed virtual electrode placement and EEG graph interpretation. The usefulness of the simulator for learning EEG measurement was tested with 35 participants randomly divided into three equal groups. Group 1 (experimental group 1) used the simulator with fuzzy feedback, group 2 (experimental group 2) used the simulator with exact feedback, and group 3 (control group) did not use a simulator. The study comprised pre- and posttests on theoretical knowledge and practical hands-on evaluation of EEG electrode placement.

Results: The Wilcoxon signed-rank test indicated that the two groups that utilized a computer-based electrode placement simulator showed significant improvement in both theoretical knowledge ($Z=1.79$, $P=.074$) and observed practical skills compared with the group that studied without a simulator.

Conclusions: Learning electrode placement using a simulator enhances students' ability to place electrodes and, in combination with practical hands-on training, increases their understanding of EEG measurement.

KEYWORDS

virtual simulation; electroencephalography; theoretical knowledge; neurophysiology; brain activity; psychomotor

XSL·FO
RenderX

Introduction

Simulation

Simulators are devices that mimic the technical and physical aspects of real-life activities and are typically used in training environments where real-life practice is not viable [1]. Simulators are as closely correlated as possible to existing guidelines and protocols of the process they are designed to imitate [2]. Room-scale simulators have multiple restrictions that mostly derive from excessive costs and size, making them unattainable by most higher education settings. For teaching and learning purposes, where multiple students would require simultaneous access to simulator training, personal computer (PC)-based simulation is seen as a viable alternative [3]. The primary function of simulators in a teaching and learning setting is not just to facilitate practical skill acquisition, but also to foster theoretical understanding [4]. Learning via PC-based simulation is experiential learning where the teacher is not always present [5]. However, students who use simulators gain better self-confidence and are less anxious when confronted with the actual event of a simulator-trained process than students who have not used simulators [6,7].

Simulations have also been used in the teaching of electroencephalography (EEG) [8]. EEG is a simple method used to monitor brain electrical waves, most commonly as a diagnostic tool for epilepsy. In EEG, electrodes are placed on the scalp and biosignals generated by cerebral neurons, modified by electrical conductivity properties of the tissues, are recorded between the electrical source and the recording electrodes. This method is widely used in neurophysiological clinical diagnosis. In EEG, evoked potentials (visual, somatosensory, motor, auditory) are examples of the recorded brain responses. In addition, different tools for diagnosing sleep disorders, such as polysomnography, respiratory polygraphy, reflex studies (blink reflex and masseter reflex), electroretinography, and electroneuromyography are examined using EEG [9].

As in all educational sectors nowadays, simulation-based approaches are increasingly being utilized in biomedical and medical education; they are used to support student learning and clinical training [10], as well as training in medical procedures [1]. In health care education, simulations are used to improve clinical performance [11] and clinical reasoning [4]. However, learning practical skills (such as EEG electrode placement) requires hands-on exercises, and in biomedical laboratory science education these exercises are usually performed in a laboratory environment [12] rather than through practical training in hospitals. Unfortunately, due to the large number of students who need EEG training, hands-on practice sessions cause disruptions at hospitals. Limited equipment and laboratory staff resources have resulted in a situation where students have a severely constrained number of EEG placement practice sessions. For this reason, teachers tend to resort to lecturing about EEG placement, which usually results in a lack of readiness of students to do practical exercises with adequate guidance.

As a result, professional EEG laboratory staff are more involved in addressing the fundamental skills (eg, identifying the head measurement starting point) of EEG electrode placement rather than providing the holistic EEG experience that a clinical laboratory session should provide. Using PC-based simulations in the field of clinical neurophysiology could be helpful to train biomedical laboratory students in the method of EEG in both practical and theoretical perspectives [13,14]. Even though PC-based simulators have become very important tools in the field of education [15], the current simulator research revealed limited academic interest and no commercially developed PC-based EEG simulator exists to date. There are some commercially available EEG estimators, but there are no educational EEG simulators for the electrode placement system. For educational purposes, some EEG simulators exist mainly for identifying EEG activity. In general, it is believed that EEG simulators are an effective, user-friendly, and inexpensive method for learning EEG morphology and recognizing seizure activity [16]. Some companies delivering EEG measurement devices have developed EEG simulators for control purposes. Efforts are currently underway to develop an EEG results simulator for ensuring that simulation quality is relevant in EEG measurement [17].

PC-Based Simulator for EEG Electrode Placement

Routine EEG examinations are implemented in clinical laboratories. Therefore, understanding the EEG recording system is very important for biomedical laboratory scientists working in clinical neurophysiology department at hospitals. In EEG, electrodes should be placed in the correct positions on a patient's scalp so that the device measurement is reliable while being as comfortable as possible for patients.

In Finland, all biomedical technicians (biomedical laboratory scientists) and nurses have graduated from universities of applied sciences with studies in clinical neurophysiology. Although the details of the study curricula across universities may differ, they all contain a theoretical (ie, classroom) component followed by practical (ie, laboratory) training. There is consensus among lecturers from several Finnish universities and laboratory professionals that the EEG method skills (ie, the underlying theory for practical work with EEG equipment) of students entering the practical training is substandard. This places an additional burden on the laboratory professionals who administer the practical training and could lead to students not being able to successfully complete their neurophysiology studies. If this situation remains unresolved, universities may soon be producing graduates who require work-under-supervision conditions or further on-the-job training before being effective and efficient biomedical technicians.

To better prepare students for practical training, we have developed a PC-based EEG simulator for learning the EEG method and neurophysiology. Our EEG simulator is based on the routinely used 10-20 mapping system [18] for electrode positioning and presents a 3-dimensional model of a human head on which students practice the placement of EEG electrodes according to this system. Our simulator relies on feedback, both immediate and summative, as its primary catalyst for learning. Feedback refers to how close to the correct position trainees place EEG electrodes on the virtual head. Instant updates, as electrodes are placed, give players an opportunity

to experience instantaneous response [19] regarding their electrode placement accuracy and an opportunity to learn by correcting their electrode positioning. In addition, our EEG simulator presents a summative scoring feedback immediately after the assignment. This satisfies student expectation because it is similar to traditional in-class learning situations where written feedback from the teacher for learning assignments and percentage grade feedback for exams are given postevent [20]. In our case, the summative feedback serves as an enabler for postsimulator debriefing [21], where students can process and strengthen their simulator learning events [22].

The EEG simulator contains two feedback systems—exact and fuzzy—for users. The fuzzy logic system provides human-like feedback through linguistic variables (ie, words) as a way to define results without a precise answer, whereas the exact system gives an axis and magnitude metric of how far away from the correct location the placement is. For example, the fuzzy feedback system provides feedback such as "placement is a little too far left" or "placement is too low," while the exact feedback system would give "placement is 6.8 mm to the left" or "placement is 32.8 mm too low." Figure 1 shows two instances of the main interactive view of the EEG simulator application, illustrating (on the right side of each image) the fuzzy and exact feedback systems. A more detailed technical explanation of the application and its implementation was presented in our previous study [8], in which students' perceptions on feedback mechanisms were examined; students initially favored the fuzzy feedback system, but after a period of practicing and improving electrode placement precision, they wanted to know the exact accuracy of their placements (ie, the exact feedback system).

The main purpose of this study was to extend our earlier research by exploring how the introduction of a PC-based EEG simulator in a higher education neurophysiology course would enhance students' acquisition of practical skills. This study was also cognizant of the importance of theoretical neurophysiology knowledge and therefore also investigated the impact of EEG simulator use on theoretical knowledge. The following hypotheses were developed for this study:

- Hypothesis 1: students utilizing the EEG simulator would show greater improvement in theoretical knowledge that those who did not study using the simulator.
- Hypothesis 2: students practicing with the EEG simulator's fuzzy feedback system would show better hands-on skills for electrode placement than those practicing with the exact feedback system and those not using the EEG simulator at all.

Figure 1. Electroencephalography simulator interface screenshots showing fuzzy (A) and exact (B) feedback systems.

Methods

Context of the Study

The context of this study was a clinical neurophysiology course that had an identical implementation of EEG training at two universities of applied sciences in southern Finland. The first author was the teacher of the training session at both universities. Figure 2 depicts a high-level comparison between the traditional course setup and our newly adapted multimodal study approach containing additional PC-based EEG simulations.

Figure 2. Traditional and adapted neurophysiology course. EEG: electroencephalography.
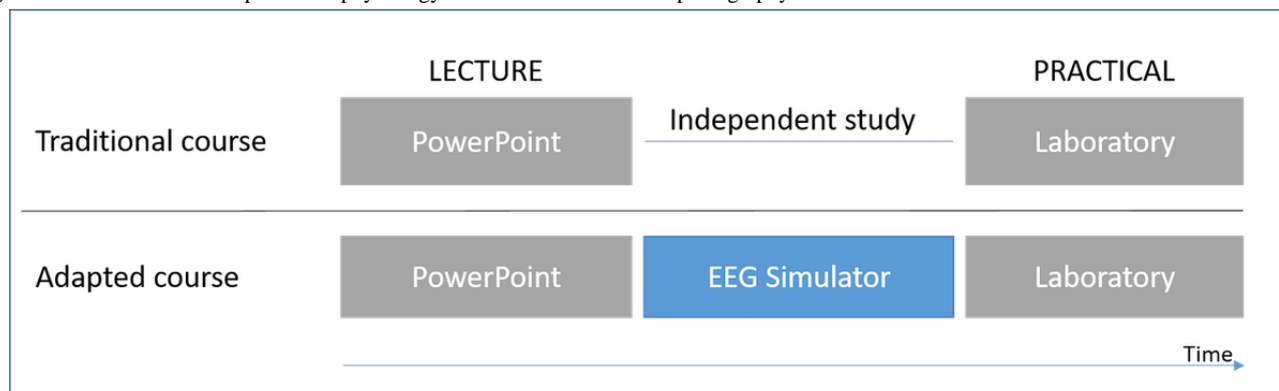
The clinical neurophysiology course included EEG and electromyography methods and evoked potentials; this study focused only on the EEG method. The course implementation and data collection were carried out in the autumn semester of 2018 at one university and in the spring semester of 2019 at the other university. Using the new multimodal study approach, both courses began with a prerecorded online lecture about the basics of the EEG method and neurophysiology. We provided supplementary study material on subject-specific issues for independent studies. All students had access to this material via an e-learning platform and were encouraged to review it at their discretion. After online lectures at the beginning of the course, students had three consecutive days to practice with the simulator for 2 hours/day under teacher guidance. Two weeks afterward, students attended the practical hands-on sessions carried out in the laboratory, where students were able to work with a real EEG hardware system to place the EEG electrodes on the scalp of a coworker's head. The final laboratory exercises were carried out at the department of clinical neurophysiology in the respective university hospitals.

Research Design and Data Collection

All students in the clinical neurophysiology course at both universities were invited to participate in the study during their respective introductory sessions, and a total of 35 students—10 male and 25 female students aged 20 to 23 years—volunteered to participate. Figure 3 shows the timeline for student activities and data collection during a 5-week data collection period. This study was piloted with 10 students before the actual study was conducted.

Figure 3. Time schedule for the study and data collections. EEG: electroencephalography; UX: user experience.
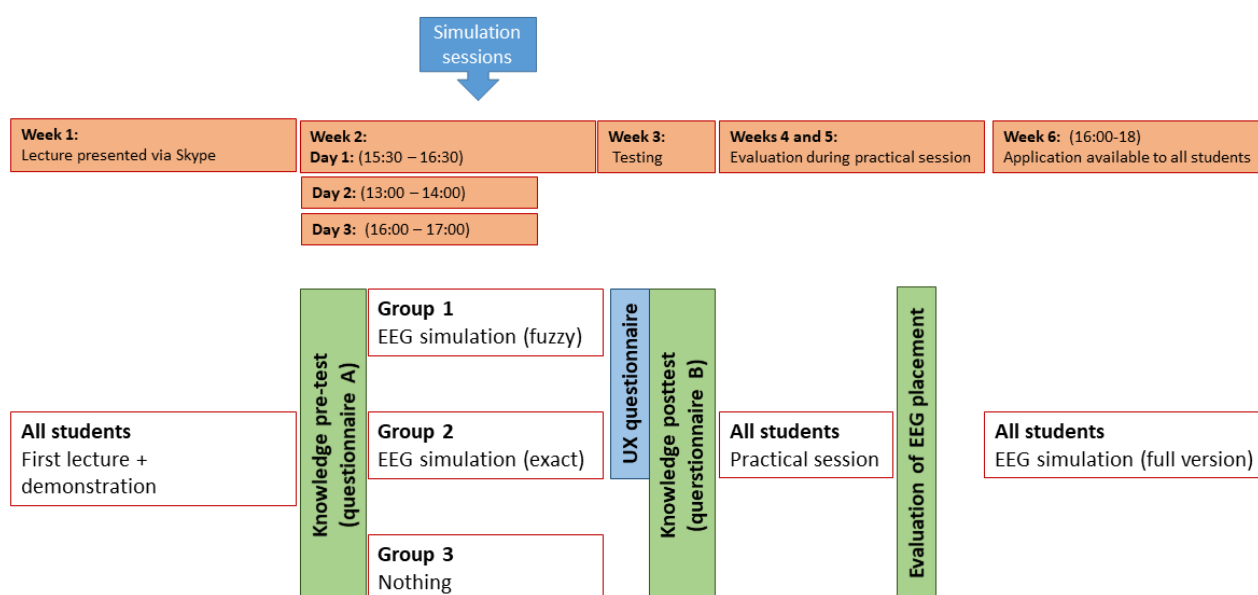
All students participated in the same prerecorded online introduction, and students' baseline theoretical knowledge was determined using a pretest (questionnaire A). After this pretest, students were randomly assigned into 3 groups: (1) group 1 studied using an EEG simulator with fuzzy feedback, (2) group 2 studied using an EEG simulator with exact feedback, and (3) group 3 (control group) studied without the EEG simulator. All groups were comprised of 11 to 12 students. After 3 guided simulation sessions for groups 1 and 2 and an independent study period for group 3, a posttest (questionnaire B) was administered to determine possible knowledge improvement. The pretest and posttest (Multimedia Appendix 1) both contained 8 questions—4 neurophysiology theory questions and 4 questions regarding the EEG method for electrode placement. The tests contained different questions to avoid the possible impact of the pretest on the posttest. To ensure that the pre- and posttest questions were valid, we selected questions from neurophysiology training guidelines. Furthermore, to safeguard that the questions were also of a suitable difficulty level, we asked the clinical teacher at the university hospital to evaluate them beforehand. During the 2-week simulation periods with the EEG simulator, students in group 3, who did not have access to the EEG simulator, were asked to study the supplementary learning materials shared on the EdX e-learning platform. Students in groups 1 and 2 also had access to these materials during our study. Students in groups 1 and 2 (who used the simulation) were instructed to keep a diary of their experiences with the simulator and asked to fill in a user experience (UX) questionnaire after using the device.

To test our hypotheses, a pretest–posttest design was implemented to characterize the effect of EEG simulations on theoretical knowledge about EEG methods and neurophysiology. The practical skill of EEG electrode placement was examined with hands-on sessions carried out in a real laboratory environment where two teachers evaluated students' EEG electrode placement skills according to EEG guidelines [10].

In order to give all students the opportunity to use the simulator, the full version of the simulator was shared with all study participants after the practical evaluations. This simulation version contained both feedback systems. In this way, each student in the study had an opportunity to learn from EEG simulation.

To help the simulator designers, UX questionnaire data were collected from everyone within 3 weeks of using the simulator.

These data provided insights to help refine the simulation process.

The usefulness of the EEG simulator and its feedback systems in practice was evaluated by observing students' accuracy on EEG placements during hands-on laboratory sessions. In addition, students were guided to keep a learning diary on their experiences and feelings during their studies.
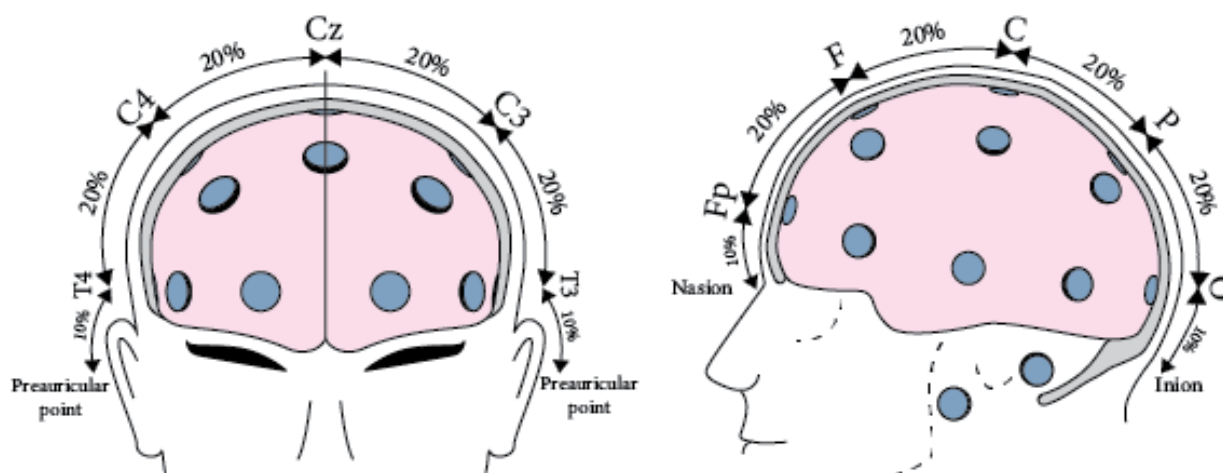
Evaluation of EEG Electrode Placement Skills

During the students' practical training of EEG electrode placement and measurement, their work was observed and evaluated. An expert clinical neurophysiologist/clinical teacher and the neurophysiology course instructor (SL and MHB) used two sets of assessment guidelines, based on the consultations with several other practicing clinical neurophysiologists. The first set focused on the students' ability to identify and measure skull dimensions, while the second set of guidelines evaluated student EEG electrode placement accuracy. The measurement assessment guidelines included the following: (1) identifying the nasion and inion points of the head, (2) measuring the distance between the nasion and inion electrode points, (3) identifying the right and left preauricular point positions on the head, and (4) measuring the head circumference.

The identification and measuring skills are important because they give the precise location of the central electrode (C_z), as well as the other electrodes' nasion-inion line. Figure 4 shows the target points from the first set of assessment guidelines on how to set up electrodes on the skull.

Figure 4. Frontal and sagittal plane views of the electroencephalography electrode placement points in the 10-20 electrode measurement system [18].

The placement assessment guidelines included the following items for evaluation: (1) accuracy of EEG electrode placement on the skull; (2) technique for measuring distance between the 3 z-points on the skull (C_z, F_z, P_z); (3) how to measure C_z electrode position on the center of the skull; (4) how to measure O_1, O_2, F_{p1}, and F_{p2} electrode positions on the skull; (5) what system was used in the EEG measurement; (6) how to identify right and left preauricular points on the head; and (7) result (cm) from the head circumference measurement.

The accurate placement of all electrodes is critical to obtaining reliable EEG data. If the electrodes are misplaced, data from an EEG scan become unrealistic and even erratic, making any diagnosis impossible. Figure 5 shows the exact placement of electrodes according to the 10-20 system.

Figure 5. Electrode positions on the skull using the 10-20 electrode electroencephalography measurement system [18].
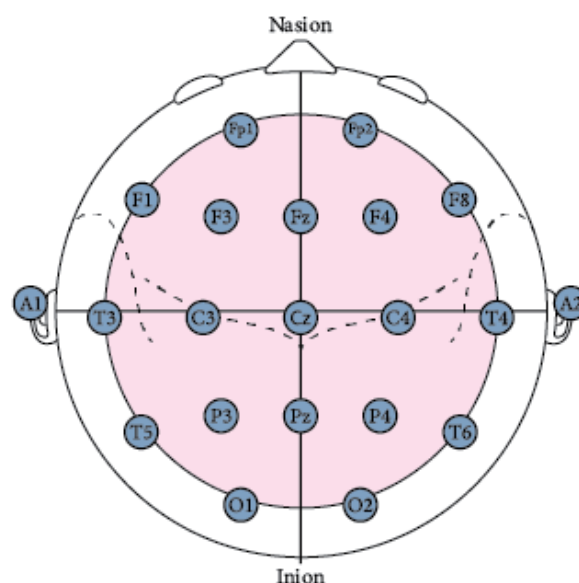
Data Analysis

Pre- and posttest answers were analyzed statistically using the SPSS software package (IBM Corp). To compare the quantitative results of pre- and posttests between groups, a Wilcoxon signed-rank test was performed. Learning diaries were analyzed by content analysis using ATLAS.Ti (ATLAS.ti Scientific Software Development GmbH) and a deductive reasoning approach. These deductions allowed the clarification of students' written answers. Practical part observations were carried out using guidelines that were created for all neurophysiology laboratories using EEG measuring methods [23].

Results

This study was comprised of 35 students. Table 1 shows the results from the pre- and posttests. Students who used the exact feedback mode in the EEG simulation scored slightly higher in their knowledge of the EEG method after the simulation than those who used the fuzzy feedback mode. Students who studied using the EEG simulator exhibited greater knowledge of the EEG method than students who did not use simulation. Only students in group 3 (students who did not use the EEG simulation) showed a significant improvement in their knowledge of neurophysiology ($Z=1.79$, $P=.074$) within the group. Group 3 also had the lowest baseline test scores of knowledge of neurophysiology. Groups 1 and 2 did not show significant improvement in their knowledge of neurophysiology after using the EEG simulator.

Table 1. Students' pre- and posttest knowledge scores within the study.

	Students' knowledge test scores					
	Group 1 (fuzzy feedback)		Group 2 (exact feedback)		Group 3 (no simulation)	
Subject area and test	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median
Neurophysiology						
Pretest (questionnaire A)	11.30 (1.70)	11.50	11.58 (1.62)	11.50	10.62 (1.85)	10.00
Posttest (questionnaire B)	12.20 (1.99)	13.00	12.80 (2.04)	13.00	13.17 (1.83)	13.00
EEG[a] method						
Pretest (questionnaire A)	9.20 (1.55)	10.00	8.17 (2.69)	8.00	9.62 (1.98)	10.00
Posttest (questionnaire B)	11.80 (2.66)	12.50	12.40 (2.27)	12.50	10.17 (2.32)	10.50

[a]EEG: electroencephalography.

At the beginning of the study, after an online prelecture and based on the results of the pretest (questionnaire A), there was no significant difference in the students' basic knowledge of the EEG method between the groups (Table 1). When comparing the mean test scores for the EEG method, a significant improvement in knowledge among students of all groups between pre- and posttests was found ($Z=3.02$, $P=.003$). When knowledge of the EEG method within the groups was compared, students who used the exact feedback system in the EEG simulation showed the greatest increase during the study

XSL•FO
RenderX

(Z=2.32, P=.021). In addition, students who used the fuzzy feedback system in the EEG simulation showed significantly higher knowledge of the EEG method in the posttest compared with the pretest (Z=2.20, P=.028). Only students who did not use the EEG simulation in their studies failed to show a significant improvement in their scores of knowledge of the EEG method from the pretest to the posttest (P=.892). A similar trend was seen in the median test scores, where the scores improved 2.50, 4.50, and 0.50 between the pre- and posttest for the fuzzy feedback group, exact feedback group, and no simulation group, respectively. Students who studied using the EEG simulation knew better how and where the EEG electrodes should be placed on the skull. However, it was also clear that these students still needed teacher support in learning the correct placements of the EEG electrodes even after the simulation.

In addition, students who used the EEG simulator were asked to write about the simulation experience—including benefits, disadvantages, motivations, and novelty—and their diary entries suggested that the EEG simulation influenced their learning, as shown in the following excerpts:

> *The simulation was helpful in learning the right electrode positions.* [S11]

> *The EEG simulation made it easier to visualize the positions of the electrodes. It explained a lot about how to do it in practice. The simulation did not help outline other things about neurophysiology.* [S1]

> *Learning motivation is enhanced by the fact that the instructions, for example with the help of pictures, are clear. The trial-and-error approach to learning that the simulator brings, however, does not necessarily increase motivation.* [S12]

Based on the learning diaries, 45.7% (16/35) of the students indicated that PC-based simulations are generally useful in education. In addition, one-half (18/35, 51.4%) of the students indicated that they would be ready to use their own time to practice the EEG method and electrode placements using a PC-based simulator. However, several students mentioned that hands-on skill practice cannot be totally replaced by virtual simulation, as those skills need to be practiced in the laboratory environment as well.

> *I was positively surprised by the introduction of an EEG placement simulator, and the initial idea that this could provide some help with learning. Of course, simulations are no substitute for practice.* [S1]

> *The electrode placement during simulator utilization sessions was continuously improving and I was getting faster at placing electrodes every time I restarted. The simulation provides a good foundation before the actual hands-on practical electrode placement session. However, there is very little theory in the simulation.* [S11]

Unfortunately, some absenteeism in the evaluation of practical skills brought small changes in the number of participants. Eight students from group 1 (fuzzy feedback), 12 students from group 2 (exact feedback), and 12 students from group 3 (no simulation) completed the practical EEG electrode placement session. Since placement accuracy is vital to EEG diagnosis, students were evaluated as being "successful" or "unsuccessful" in their task. Only students who placed all electrodes accurately were considered to have successfully completed the task. Table 2 presents the number of students from each group who successfully or unsuccessfully completed the electrode placement task.

Table 2. Results from practical evaluations of electrode placement.

Measurement	Group 1 (fuzzy feedback; n=8)		Group 2 (exact feedback; n=12)		Group 3 (no simulation; n=12)	
	Successful	Unsuccessful	Successful	Unsuccessful	Successful	Unsuccessful
C_z electrode position	8	0	12	0	9	3
O_1, O_2, F_{p1}, and F_{p2} electrode positions	8	0	10	2	9	3
10-20 EEG[a] system	8	0	12	0	6	6
Right and left preauricular points	8	0	12	0	9	3
Head circumference	8	0	11	1	7	5

[a]EEG: electroencephalography.

Evaluation through observation of the practical placements confirmed the results of the pre- and posttests. Students who studied using the EEG simulator showed greater skills at placing the electrodes. When comparing the results of groups 1 (fuzzy feedback) and 2 (exact feedback), students who used simulation with the fuzzy (human-friendly) feedback system performed the best in practical placements. These students demonstrated a superior placement accuracy compared with the students from group 2, the exact feedback system, who had problems in O_1, O_2, F_{p1}, and F_{p2} electrode position measurements, as well as in

head circumference measurement. Students who did not use the simulator (group 3) had difficulty overall in the EEG electrode placement. They managed to complete the practical part of the task (ie, the EEG electrode placement) only with the support of the teacher.

Some students in group 3 (no simulation group) expressed that placing the EEG electrodes on their coworker's head was challenging because electrodes were moving on the scalp. Some students in group 2 (exact feedback group) had problems getting O_1, O_2, F_{p1}, and F_{p2} electrodes to work properly. However, this

XSL•FO

RenderX

was a consequence of an inexact placement of EEG electrodes. Students using the fuzzy feedback system (group 1) indicated that the EEG simulation helped them to better remember the placement of electrodes and distances from each other on the head. On the other hand, these students commented that studying by simulation was not the same as placing the electrodes in a practical setting because the simulation did not include hands-on work, as seen in the following excerpt:

> *The simulation helped, but it was good to see and experience how the electrode placement was actually done… which tools are used and how to measure them. Electrodes placement vary from laboratory to laboratory.* [S16]

Discussion

Principal Findings

The results confirm the hypothesis that studying using a simulator provides additional support for learning the EEG method and showed a positive influence in students' learning of neurophysiology. The increase in knowledge may at least partly be a consequence of increased motivation of those students that used the simulation during their studies. The results indicate that simulation with a logical (human-friendly fuzzy feedback) system has a more positive impact on practical skills, but the exact feedback simulation is an important tool from the theoretical knowledge development point of view. On the other hand, the theoretical knowledge of those students who did not use simulation increased the most, especially concerning the basics of neurophysiology, thereby refuting our first null hypothesis. This can be explained by the fact that those students who did not use the simulator may have had more time to concentrate and to study the theoretical supplementary materials provided in the course. We noted that the diary entries of the students using the simulator did not mention any additional engagement with theory learning materials to augment their practice of the EEG method. It may also be that because the nonsimulator students knew that they were part of a study that was measuring learning, they made extra efforts to learn the material. However, these students were not able to achieve all additional learning that the exact and fuzzy simulator groups enjoyed. Learning hands-on skills requires some instruments, and for the EEG method, this learning was enhanced with PC-based simulation.

This study revealed that for students using the EEG simulator, their knowledge of the EEG method increased more than their knowledge of neurophysiology. The insignificant improvement in their knowledge of neurophysiology through use of EEG simulation was also due to the fact that the simulator did not require the students to be immersed in theoretical physiology-related material. This is in line with the results of Jaakkola et al [12], which suggest that the best learning results are gained using a multimodal learning approach, where virtual simulations are done together with practical exercises. Nevertheless, the modest improvement in neurophysiology knowledge among students using the simulator resulted from the requirement to learn the 10-20 system, which links to some aspects of neurophysiology theory.

Although students preferred the fuzzy feedback mode, the purposeful nature of their encounters with the learning material to test our second hypothesis explains why both the fuzzy and exact feedback systems led to significant improvement in knowledge of the EEG method. Students who did not use the simulator did not show significant improvement in their knowledge of the EEG method simply because they had no means of developing a practical frame of reference for electrode placement. This result is in line with that of Miller et al [16], who used a different EEG simulator, albeit without an electrode placement feature.

A further study by Bottomley et al [24] indicated that visualization increases the enjoyment of learning, making students more susceptible to learning. Although we did not study learning strategies, this observation of visual learning cannot be considered a coincidence, but it requires further study.

The learning diaries were only written during the simulation period. In other words, the students had no exposure to practical sessions at the time of keeping their diaries and the diaries were returned before the practical sessions. Students realized early on that a PC-based simulator could supplement hands-on training but could not replace it. It is clear that PC-based simulations cannot replace teachers themselves, as teachers are needed to familiarize students with new virtual learning environments. Improved technology only provides new tools for education but does not have an effect on the studied learning content [25]. Students' comments in diaries indicated that less than half of the students believed that the EEG simulator might be useful, but the results clearly indicated that the simulator significantly influenced their learning. Furthermore, half of the students would be willing to invest their own time into using the simulator, meaning that the other half would require further encouragement, for instance through continued teacher oversight and guidance.

The students who used the simulator were more pragmatically engaged with the learning material throughout and therefore more inclined to remember all the steps when conducting the practical hands-on task. These research results support the notion that the use of PC-based simulations in education should be supported and encouraged.

Conclusions and Future Work

In this study, we used an experimental research design to understand what influence a PC-based EEG simulator could have on both theoretical knowledge and practical skills among higher education biomedical laboratory science students. By measuring the learning outcomes, we were able to gauge whether the introduction of a PC-based simulator could better prepare students for practical hands-on sessions.

Although our study raised some reliability concerns due to a limited number of participants across 3 experimental groups, we remain confident that utilizing a PC-based EEG csimulator in neurophysiology classes gives students more opportunities and increases motivation to learn practical EEG placement. It did not, however, appear to improve learning of neurophysiology theory. It was also clear that students recognized the value of hands-on work when it comes to learning practical skills.

Students strongly believed that the actual handling of EEG equipment is invaluable in learning how to use it, leading us to recommend that simulator training should supplement hands-on training and never attempt to replace it. Using a teaching strategy that complements theory lessons with PC-based simulator practice proved to significantly enhance practical learning of the EEG method among higher education biomedical laboratory science students, making students better prepared for hands-on training. We hereby postulate that using PC-based EEG simulators in courses that include other electrical medical devices, such as respirators and electrocardiographs, will also improve practical application knowledge of these devices.

This study completed a second design-develop-test-evaluate iterative cycle of an ongoing endeavor to improve neurophysiology competence among university students in the applied sciences. Future work will include minor revisions to some of the EEG simulator features and a longitudinal study to examine the long-term learning effects of using our simulator. More immediate efforts have gone into developing a mobile augmented reality application to complement the neurophysiology theory learning material. We are exploring whether such an application would stimulate enough intrinsic motivation to engage with the theoretical part of the course, thereby also improving conceptual understanding.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Pretest (questionnaire A) and posttest (questionnaire B).
[DOCX File , 13 KB - games_v8i4e18768_app1.docx]

References

1. Shanks D, Wong RY, Roberts JM, Nair P, Ma IW. Use of simulator-based medical procedural curriculum: the learner's perspectives. BMC Med Educ 2010 Nov 08;10:77 [FREE Full text] [doi: 10.1186/1472-6920-10-77] [Medline: 21059253]
2. Langergraber G, Rieger L, Winkler S, Alex J, Wiese J, Owerdieck C, Ahnert M, Simon J, Maurer M. A guideline for simulation studies of wastewater treatment plants. Water Sci Technol 2004;50(7):131-138. [doi: 10.2166/wst.2004.0436] [Medline: 15553468]
3. Evgueni Khvilon K. Information and communication technologies in teacher education: a planning guide. Unesco. 2002. [WebCite Cache ID https://unesdoc.unesco.org/ark:/48223/pf0000129533]
4. Cleave-Hogg D, Morgan PJ. Experiential learning in an anaesthesia simulation centre: analysis of students' comments. Med Teach 2002 Jan 01;24(1):23-26. [doi: 10.1080/00034980120103432] [Medline: 12098453]
5. Ivancic K, Hesketh B. Learning from errors in a driving simulation: effects on driving skill and self-confidence. Ergonomics 2010 Nov 10;43(12):1966-1984. [doi: 10.1080/00140130050201427]
6. Murray C, Grant MJ, Howarth ML, Leigh J. The use of simulation as a teaching and learning approach to support practice learning. Nurse Education in Practice 2008 Jan;8(1):5-8. [doi: 10.1016/j.nepr.2007.08.001]
7. Jaakkola T, Nurmi S, Veermans K. A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts. J. Res. Sci. Teach 2010 Apr 27;48(1):71-93. [doi: 10.1002/tea.20386]
8. Björn M, Ravyse W, SanzVillafruella D, Luimula M, Leivo S. Higher Education Learner Experience with Fuzzy Feedback in a Digital Learning Environment. 2018 Presented at: IEEE International Conference on Cognitive Infocommunications (CogInfoCom); 22-24 August 2018; Budapest, Hungary p. 22-24. [doi: 10.1109/coginfocom.2018.8639910]
9. Kliininen neurofysiologia. Suomen Bioanalyytikkoliitto ry, Finlands Bioanalytikerförbund rf [Association of Biomedical Laboratory Scientists in Finland]. URL: https://www.bioanalyytikkoliitto.fi/mika-ihmeen-bioanalyytikko/bioanalyytikon-koulutus/erikoisalat/kliininen-neurofysiologia/
10. Al-Elq A. Simulation-based medical teaching and learning. J Family Community Med 2010 Jan;17(1):35-40 [FREE Full text] [doi: 10.4103/1319-1683.68787] [Medline: 22022669]
11. Liaw SY, Wong LF, Chan SW, Ho JTY, Mordiffi SZ, Ang SBL, et al. Designing and Evaluating an Interactive Multimedia Web-Based Simulation for Developing Nurses' Competencies in Acute Nursing Care: Randomized Controlled Trial. J Med Internet Res 2015 Jan 12;17(1):e5. [doi: 10.2196/jmir.3853]
12. Jaakkola T, Nurmi S, Veermans K. A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts. J. Res. Sci. Teach 2010 Apr 27;48(1):71-93. [doi: 10.1002/tea.20386]
13. Borgar E. Knowledge fields and inner patterns in clinical laboratory science. Clin Lab Sci 2002;15(2):102-10, 115. [Medline: 12776773]

XSL•FO

RenderX

14. Olejniczak P. Neurophysiologic basis of EEG. J Clin Neurophysiol 2006 Jun;23(3):186-189. [doi:
     10.1097/01.wnp.0000220079.61973.6c] [Medline: 16751718]

15. Chini JJ, Straub CL, Thomas KH. Learning from avatars: Learning assistants practice physics pedagogy in a classroom
     simulator. Phys. Rev. Phys. Educ. Res 2016 Feb 22;12(1). [doi: 10.1103/PhysRevPhysEducRes.12.010117]

16. Miller B, Troyer M, Busey T. Virtual EEG: A Software-Based Electroencephalogram Designed for Undergraduate
     Neuroscience-Related Courses. J Undergrad Neurosci Educ 2008;7(1):A19-A25 [FREE Full text] [Medline: 23493937]

17. Kemp B, Zwinderman A, Tuk B, Kamphuisen H, Oberyé J. Analysis of a Sleep-Dependent Neuronal Feedback Loop: The
     Slow-Wave Microcontinuity of the EEG. September, Vol. 2000. URL: https://helka.finna.fi/Primo/Search?lookfor=
     %22IEEE+Transactions+on+Biomedical+Engineering%22&type=AllFields [accessed 2000-09-01]

18. Klem G, Lüders H. The ten-twenty electrode system of the International Federation. Recommendations for the Practice of
     Clinical Neurophysiology: Guidelines of the International Federation of Clinical Physiology (EEG Suppl. 52, Editors: G.
     DeutschlA. Eisen International Federation of Clinical Neurophysiology. Published by Elsevier Science B. V 1999. [doi:
     10.1016/s1567-424x(10)61018-5]

19. Cheng M, Lin Y, She H. Learning through playing Virtual Age: Exploring the interactions among student concept learning,
     gaming performance, in-game behaviors, and the use of in-game characters. Computers & Education 2015 Aug;86:18-29.
     [doi: 10.1016/j.compedu.2015.03.007]

20. Kim J, Na BJ, Yun J, Kang J, Han S, Hwang W, et al. What kind of feedback do medical students want? Korean J Med
     Educ 2014 Sep 01;26(3):231-234. [doi: 10.3946/kjme.2014.26.3.231]

21. Baranowski T, Baranowski J, Thompson D, Buday R, Jago R, Griffith MJ, et al. Video Game Play, Child Diet, and Physical
     Activity Behavior Change. American Journal of Preventive Medicine 2011 Jan;40(1):33-38. [doi:
     10.1016/j.amepre.2010.09.029]

22. Crookall D. Engaging (in) Gameplay and (in) Debriefing. Simulation & Gaming 2015 Jan 15;45(4-5):416-427 [FREE Full
     text] [doi: 10.1177/1046878114559879]

23. Acharya JN, Hani A, Cheek J, Thirumala P, Tsuchida TN. American Clinical Neurophysiology Society Guideline 2. Journal
     of Clinical Neurophysiology 2016;33(4):308-311. [doi: 10.1097/wnp.0000000000000316]

24. Bottomley S, Chandler D, Morgan E, Helmerhorst E. jAMVLE, A new integrated molecular visualization learning
     environment. Biochem Mol Biol Educ 2006 Sep;34(5):343-349 [FREE Full text] [doi: 10.1002/bmb.2006.494034052666]
     [Medline: 21638712]

25. Liikanen E, Björn M, Nielsen M. Use of information and communications technology by teachers and students in biomedical
     laboratory science educations in the Nordic countries. Educ Inf Technol 2018 Jun 9;23(6):2867-2878. [doi:
     10.1007/s10639-018-9746-y]

Abbreviations

EEG: electroencephalography
PC: personal computer
UX: user experience

XSL•FO
RenderX

<u>Original Paper</u>

Effects of Virtual Reality and Non–Virtual Reality Exercises on the Exercise Capacity and Concentration of Users in a Ski Exergame: Comparative Study

Junho Ko[1], PhD; Seong-Wook Jang[1], PhD; Hyo Taek Lee[2], PhD; Han-Kyung Yun[3], PhD; Yoon Sang Kim[1], PhD

[1]BioComputing Lab, Institute for Bio-engineering Application Technology, School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan-si, Republic of Korea

[2]Department of Sport Interaction Science, Sungkyunkwan University, Suwon-si, Republic of Korea

[3]School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan-si, Republic of Korea

Corresponding Author:
Yoon Sang Kim, PhD
BioComputing Lab, Institute for Bio-engineering Application Technology, School of Computer Science and Engineering, Korea University of Technology and Education
1600, Chungjeol-ro, Byeongcheon-myeon, Dongnam-gu
Cheonan-si, 31253
Republic of Korea
Phone: 82 24273295
Email: yoonsang@koreatech.ac.kr

Abstract

Background: Recently, ski exergames have been gaining popularity due to the growing interest in health improvement. Conventional studies evaluating the effects of ski exergames only considered exercise capacity and overlooked concentration. Ski exergames consist of a motion platform for exercise and virtual reality (VR) content in the game. The VR content enhances the exercise capacity and concentration of the user by providing a challenging goal.

Objective: The aim of this study is to evaluate the effects of VR and non-VR exercises on the exercise capacity and concentration of users in a ski exergame.

Methods: To examine the effects of the VR content in ski exergames, we performed 2 experiments, non-VR exercise and VR exercise, where participants exercised on the motion platform. If a user performs an exercise without using any VR content, it is a non-VR exercise. Contrastingly, in the case of VR exercise, a user exercises according to the VR content (a downhill scenario). In addition to the range of motion (ROM) of the ankle and rated perceived exertion (RPE) to assess exercise capacity, we used electroencephalography (EEG) to assess users' concentration.

Results: We evaluated the effects of the VR content by comparing the results obtained from VR and non-VR exercises. The ROM of the ankle with VR exercise was wider than that with non-VR exercise. Specifically, ROM of the ankle was 115.71° (SD 17.71°) and 78.50° (SD 20.43°) in VR exercise and non-VR exercise, respectively. The RPE difference between the 2 exercises was not statistically significant. The result of the sensorimotor rhythm waves (which are concentration-related EEG signals) was more favorable for VR exercise than non-VR exercise. The ratios of sensorimotor rhythm wave in EEG were 3.08% and 2.70% in the VR exercise and non-VR exercise, respectively.

Conclusions: According to the results of this experiment, higher exercise capability and concentration were achieved with the VR exercise compared with non-VR exercise. The observations confirm that VR content can enhance both exercise capability and concentration of the user. Thus, the ski exergames can be used effectively by those who, in general, do not like exercise but enjoy games.

XSL•FO
RenderX

Introduction

Exergames (a portmanteau of "exercise" and "games") are interactive video games that stimulate an active, whole-body gaming experience [1]. Exergames are gaining popularity, mainly, due to the growing interest in health improvement [2]. Depending on the purpose, exergames are classified into 3 categories, namely, exergames for (1) elite athletes' training, (2) rehabilitation of patients, and (3) health improvement of general people [3]. Due to their requirements of large space and high costs, exergames have been traditionally used to train elite athletes or rehabilitate patients. However, recent years have witnessed cost reduction and miniaturization of exergames using technologies such as micro-electromechanical systems, which allowed the dissemination of various exergames to improve health of the general public [4].

Generally, high levels of an aerobic exercise are reported to enhance the concentration as well as exercise capability [5]. In addition, a fast video game has been reported to enhance concentration by requiring real-time monitoring for large information [6]. Conventional studies have suggested that an exergame, combined with high levels of aerobic exercise and a fast video gaming experience, enhances exercise capability and concentration [1]. However, a few exergames have been studied for their effect on the concentration of users [7]. Therefore, we would like to evaluate the effect of ski exergames with high level of aerobic exercise on the general public.

Although skiing is a popular sport, it is limited by environmental seasons and location. To address these limitations, ski simulators for indoor training, which can be used anytime and anywhere, have been developed [8-10]. Recently, the commercialization of ski simulators [11-13] has been useful in not only training elite athletes but also improving people's health.

Ski exergames consist of a motion platform and virtual reality (VR) content. The motion platform enables users with the natural skiing experience through feedback (output) according to the input motion. The VR content provides the user with a virtual ski environment (with multimodalities such as vision and sound). The VR content suggests a challenging goal for motivation (competition and achievement), and the suggested goal improves the exercise capacity and concentration of the user [14]. In the ski exergame, concentration for finding a downhill route and exercise capacity for maintaining a posture are required, such as the actual ski motion. However, the conventional studies [15-17] evaluating the effect of ski exergames considered only the exercise capacity, not concentration. Therefore, this study investigates the effects of VR content–enabled ski exergame on the exercise capacity and concentration of the user. This study compares the effects of VR and non-VR exercise on the exercise capacity and concentration of users in the ski exergame.

Methods

Design and Setting

This study was approved by the institutional review board (IRB-2018-2-10) of Korea University of Technology and Education (KOREATECH), South Korea. To examine the effect of the VR content in the ski exergames, we performed 2 experiments in which participants exercised on a motion platform with and without VR content. We measured the range of motion (ROM) of the ankle and rated perceived exertion (RPE) of the ankle to assess exercise capacity. Moreover, electroencephalography (EEG) was performed to assess concentration of the user. An experimental environment is shown in Figure 1.

Figure 1. Experimental environment. EEG: electroencephalogram; IMU: inertial measurement unit; ROM: range of motion; VR: virtual reality.
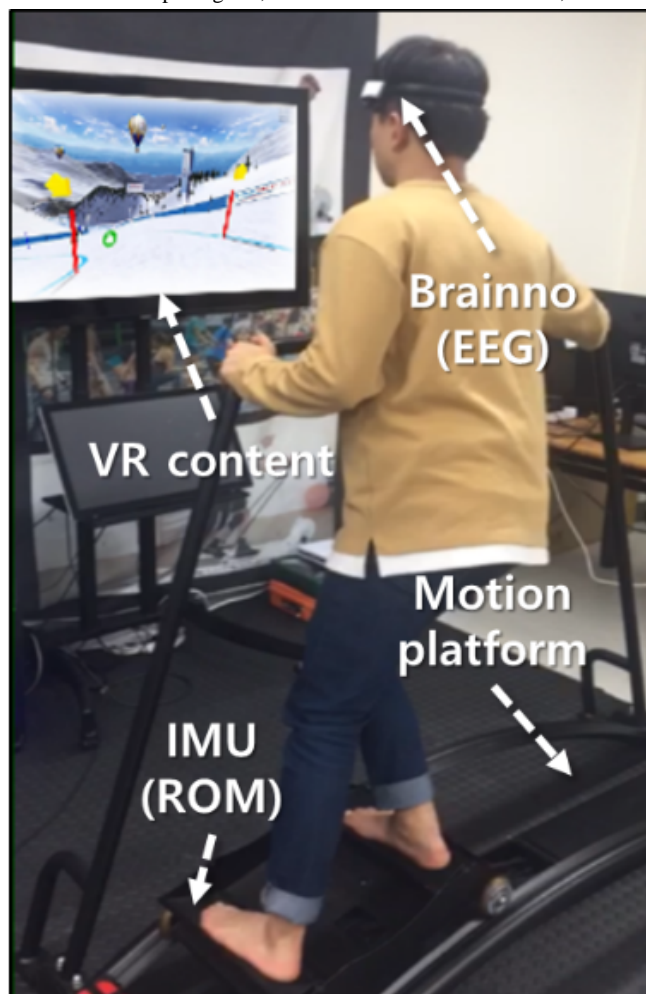
The ski motion platform (Basic Ski Simulator, Pro Ski Simulator) [11] enables the users to experience real skiing by moving the foothold side to side. Using the VR content (Ski Fit 360, Studio 360 Connect) [12], users can experience alpine ski racing (a virtual downhill scenario, where a user passes through 55 gates in a minute). Here, the ROM of the ankle in the sagittal plane is measured using an inertial measurement unit (IMU). Brainno [18] is a device that measures the EEG noninvasively using 2-channel dry type electrodes. It records the measured EEG according to frequency bandwidth of delta (0~3 Hz), theta (4~7 Hz), alpha (8~12 Hz), sensorimotor rhythm (SMR, 13~15 Hz), M-beta (16~20 Hz), H-beta (21~30 Hz), and gamma (over 31 Hz) waves.

Participants

Participants in the experiment were 10 adults with no physical disabilities. The participants having neurosurgery and cardiovascular history were excluded in the subject selection process. Table 1 shows participant demographics, namely, age, height, weight, BMI, and percent body fat.

Table 1. Participant demographics (N=10).

Characteristics	Values, mean (SD)
Age (years)	28.50 (6.11)
Height (cm)	173.24 (8.64)
Weight (kg)	69.88 (17.55)
BMI (kg/m^2)	23.17 (4.90)
Body fat (%)	26.77 (7.18)

The subjects were provided with all the essential explanations for the purpose and method of the experiments before the 2 experiments were conducted. The subjects agreed to participate in the experiment.

Procedures

The exercises were divided into 2 categories: non-VR exercise (without VR content) and VR exercise (with VR content). In the non-VR exercise, the subject had to exercise freely without the VR content, while in the VR exercise, the subject had to

exercise according to the VR content (downhill scenario). The order of the experiments (non-VR exercise and VR exercise) for the participants was randomly assigned. To minimize the effect of exertion on the next experiment, we assigned a rest period of 30 minutes between 2 exercises. The exercise procedures for both the cases consist of stretching, rest, exercise, and RPE survey, as listed in Table 2.

Table 2. Experimental procedure.

	Experiment 1 (Non-VR[a] exercise without VR content)				Rest	Experiment 2 (VR exercise with VR content)			
Exercise procedures	Stretching	Rest	Exercise	RPE[b] survey		Stretching	Rest	Exercise	RPE survey
Time (minutes)	10	2	1	2	30	10	2	1	2

[a]VR: virtual reality.

[b]RPE: rated perceived exertion.

In each experiment, the participants freely stretched for 10 minutes and took rest for 2 minutes before the exercise. Subsequently, the participants exercised for 1 minute, followed by RPE recording based on the modified Borg scale (the score was given on the scale of 0-10).

Both ankle ROM and EEG data of the subjects were collected during exercise. To obtain the ROM of the ankle, the IMU measured the rotation angle including inversion and eversion. The rotation angle was based on an axis of the IMU parallel to the normal of the sagittal plane. To obtain the EEG, the Brainno device measured brain waves as time-series data at 256 Hz sampling rate. The time-series data were recorded as a rate (%) of EEG activation through power spectrum analysis in the frequency domain; this makes it possible to quantitatively understand the weight of each component in the EEG. Artifacts from the EEG were removed using a finite impulse response filter.

Data Analysis

We analyzed the data measured from the experiments using SPSS 21 (IBM Corp). The data measured from the experiments are ROM of the ankle, RPE, and EEG. Since the measured data did not meet the normality assumptions, and the sample size was small, we applied the Wilcoxon nonparametric test to examine significant differences between the measured values. A significance level of 0.05 (95%) was used for the ROM of the ankle, RPE, and EEG analysis.

Results

The ROM of the ankle, as measured in the experiments, is shown in Figure 2. ROM of the ankle is the maximum range of rotation of the ankle joint in the sagittal plane in experiment 1 and experiment 2. The ROM of the ankle was 78.50° (SD 20.43°) and 115.71° (SD 17.71°) in experiment 1 and experiment 2, respectively, which implies that the ROM of the ankle in experiment 2 was wider than that in the non-VR exercise (P=.015).

Figure 2. ROM of ankle measured in experiment 1 and experiment 2. VR: virtual reality.
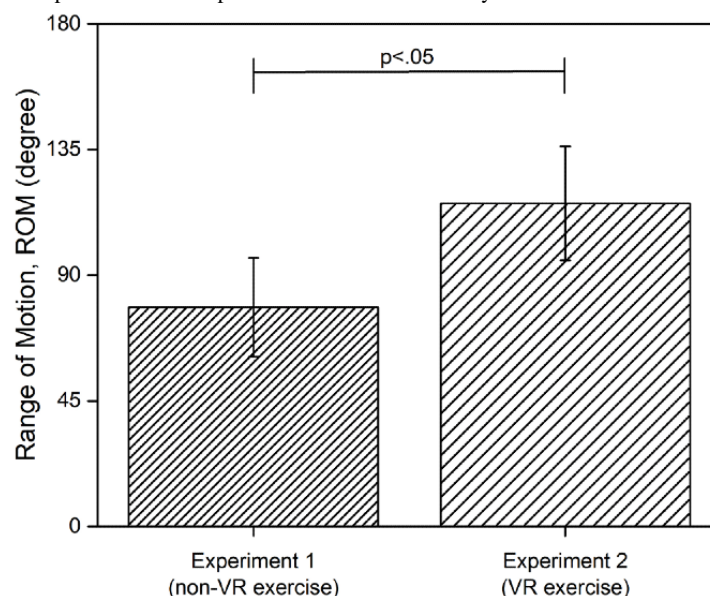
RPEs for both the experiments are shown in Figure 3. RPE is the exercise intensity measured using a survey after experiment 1 and experiment 2. The RPEs were 4.10 (SD 1.85) and 2.20 (SD 0.42) for experiment 1 and experiment 2, respectively.

However, we did not find any statistically significant difference between the RPEs of the 2 exercises (the non-VR exercise and VR exercise).

The EEG data measured from the experiments are listed in Table 3.

Figure 3. The RPE measured in experiment 1 and experiment 2. VR: virtual reality.
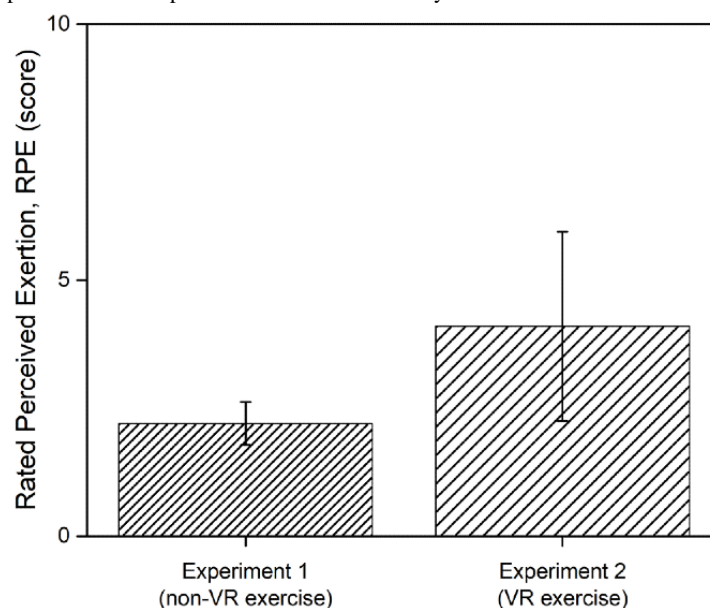

Table 3. EEG measurement in experiment 1 and experiment 2.

EEG[a]	Experiment 1 (non-VR[b] exercise without VR content; %)	Experiment 2 (VR exercise with VR content; %)	Wilcoxon *P* value
Delta	55.26	53.10	.226
Theta	16.73	15.70	.705
Alpha	7.89	8.35	.290
SMR[c]	2.70	3.08	.017[d]
M-Beta	3.63	4.05	.140
H-Beta	5.95	6.87	.545
Gamma	7.84	8.84	.364
Total	100	100	.705

[a]EEG: electroencephalography.

[b]VR: virtual reality.

[c]SMR: sensorimotor rhythm.

[d]Significant differences (P<.05)

EEG is the ratio of brain waves measured from experiment 1 and experiment 2. We did not find any significant differences with respect to the Delta, Theta, Alpha, M-Beta, H-Beta, and Gamma waves. However, significant differences were observed with respect to the SMR wave (*P*=.017) between non-VR exercise and VR exercise. The ratios of sensorimotor rhythm wave in EEG were 2.70% and 3.08% in experiment 1 and experiment 2, respectively. The SMR wave is related to the concentration, and the corresponding results are favorable for VR exercise over non-VR exercise.

Discussion

In this study, we investigated the effects of the VR and non-VR exercises on the exercise capacity and concentration of users in the ski exergame. To provide resistance between snow surface and ski plate, the ski motion platform was fitted with an elastic band that stretches linearly with the tension from the start of contraction to the maximum ROM of the ankle. Due to such mechanical characteristics of the system, it is hard for a user on the ski motion platform to maintain a balanced posture if the user is moving from side to side. Moreover, since it is harder to maintain the posture, the user has to significantly move the lower limb joints, including the ankle [15]. In this experiment, we found that the ROM of the ankle in the VR exercise was wider than that in the non-VR exercise. This implies that the VR content in the VR exercise induced a wider side-to-side movement (exercise) than the non-VR exercise, and the users rotated the ankle joint to larger angles to maintain the posture.

The SMR wave in an EEG occurs in the state of attention and activity. This type of wave is mainly observed if the subject is solving the problem that requires a simple concentration [19]. From the results of this experiment, we found that the SMR wave favors VR exercise more than non-VR exercise, which

implies that VR content used in the VR exercise enhanced the concentration more than that in the non-VR exercise.

We found that VR content enhances challenge, motivation, and concentration by setting aims of higher gate-passing accuracy and reduced racing time. Our study confirmed that higher concentration could be achieved with VR exercise using VR content than with non-VR exercise. Exergame is expected to be used effectively by those who, in general, do not like exercise but enjoy the game. In addition, the people who spend many hours sitting may find ski exergame useful for their health maintenance.

Since the work presented here is based on a pilot study with a small sample size, it is difficult to generalize the results. Therefore, future work should focus on increasing the number of participants and duration of the experiments. Additionally, the changes in body composition of the subjects should be measured.

As exergames persuade users to exercise and enhance their concentration through leisure, these are used by the subjects of various categories such as children, adolescents, older people, and patients [20-23]. Exergames have been reported to enhance the physical activity in low- or mid-intensity exercise regimes, similar to walking or jogging [24,25], and cognitive function related to concentration [26]. However, it is not yet known whether these effects are achieved by VR exercise rather than non-VR exercise. Many conventional studies have shown that aerobic exercise or physical activity of moderate intensity is helpful in improving the concentration (related to cognitive abilities) of children [27-30]. The results of these studies do not elucidate which of the exercises (VR or non-VR) has a larger effect on concentration and exercise capacity. Therefore, the effects of exergame should be evaluated by considering the game and exercise factors separately in future studies.

Conflicts of Interest

None declared.

References

1. Best J. Exergaming in Youth: Effects on Physical and Cognitive Health. Z Psychol 2013 Apr 01;221(2):72-78 [FREE Full text] [doi: 10.1027/2151-2604/a000137] [Medline: 25097828]
2. Miles H, Pop S, Watt S, Lawrence G, John N. A review of virtual environments for training in ball sports. 2012 Oct;36(6):714-726 [FREE Full text] [doi: 10.1016/j.cag.2012.04.007]
3. Pasch M, Bianchi-Berthouze N, van Dijk B, Nijholt A. Movement-based sports video games: Investigating motivation and gaming experience. Entertainment Computing 2009 Apr;1(2):49-61 [FREE Full text] [doi: 10.1016/j.entcom.2009.09.004]
4. Berglund A, Berglund E, Siliberto F, Prytz E. Effects of reactive and strategic game mechanics in motion-based games. : IEEE; 2017 Jun 8 Presented at: 2017 IEEE 5th International Conference on Serious Games and Applications for Health(SeGAH); 2017 Apr 2-4; Perth, WA, Australia. [doi: 10.1109/SeGAH.2017.7939275]
5. Hillman C, Pontifex MB, Raine D, Castelli DM, Hall EE, Kramer AF. The effect of acute treadmill walking on cognitive control and academic achievement in preadolescent children. Neuroscience 2009 Mar 31;159(3):1044-1054 [FREE Full text] [doi: 10.1016/j.neuroscience.2009.01.057] [Medline: 19356688]
6. Green CS, Bavelier D. Action video game modifies visual selective attention. Nature 2003 May 29;423(6939):534-537. [doi: 10.1038/nature01647] [Medline: 12774121]
7. Staiano A, Calvert S. Exergames for Physical Education Courses: Physical, Social, and Cognitive Benefits. Child Dev Perspect 2011 Jun;5(2):93-98 [FREE Full text] [doi: 10.1111/j.1750-8606.2011.00162.x] [Medline: 22563349]
8. Aleshin V, Klimenko S, Manuilov M, Melnikov L. Alpine skiing and snowboarding training system using induced virtual environment. In: Müller E, Lindinger S, Stöggl T, editors. Science and skiing IV. Aachen, Germany: Meyer & Meyer Fachverlag und Buchhandel GmbH; Jan 27, 2009:137-144.
9. Hasegawa S, Ishijima S, Kato F, Mitake H, Sato M. Realtime sonification of the center of gravity for skiing. New York, USA: Association for Computing Machinery; 2012 Mar 9 Presented at: The 3rd Augmented Human International Conference; 2012 Mar 8-9; Megève, France p. 1-4. [doi: 10.1145/2160125.2160136]
10. Nozawa T, Wu E, Koike H. VR Ski Coach: Indoor Ski Training System Visualizing Difference from Leading Skier. : IEEE; 2019 Mar 15 Presented at: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces(VR); 2019 Mar 23-27; Osaka, Japan p. 1341-1342. [doi: 10.1109/vr.2019.8797717]
11. Basic Ski-Simulator. URL: https://www.ski-simulator.com/ [accessed 2020-09-15]
12. SKI Fit 360. URL: https://www.skifit360.fr/ [accessed 2020-09-15]
13. SkyTechSport Ski Simulator. URL: https://www.skytechsport.com/ [accessed 2020-09-15]
14. Huang H, Pham T, Wong M, Chiu H, Yang Y, Teng C. How to create flow experience in exergames? Perspective of flow theory. Telematics and Informatics 2018 Aug;35(5):1288-1296 [FREE Full text] [doi: 10.1016/j.tele.2018.03.001]

15. Lee H, Kim Y, Kim Y. Kinematic study with and without ski boots using ski simulator. Science & Sports 2017 Feb;32(1):e9-e14 [FREE Full text] [doi: 10.1016/j.scispo.2016.07.008]

16. Lee H, Kim Y, Kim Y. Effects of unilateral, bilateral movement and combined exercise using a ski simulator on lower limb joint and muscle activities. Science & Sports 2018 Jun;33(3):176-182 [FREE Full text] [doi: 10.1016/j.scispo.2017.11.003]

17. Lee H, Kim Y, Roh H. Changes in the Lower Limb Joint Angle during the Simulated Skiing. J Phys Ther Sci 2012;24(6):471-474 [FREE Full text] [doi: 10.1589/jpts.24.471]

18. Brainno. URL: http://soso-g.co.kr/ [accessed 2020-09-15]

19. Sterman MT, Egner T. Foundation and practice of neurofeedback for the treatment of epilepsy. Appl Psychophysiol Biofeedback 2006 Mar;31(1):21-35 [FREE Full text] [doi: 10.1007/s10484-006-9002-x] [Medline: 16614940]

20. Edwards J, Jeffrey S, May T, Rinehart NJ, Barnett LM. Does playing a sports active video game improve object control skills of children with autism spectrum disorder? J Sport Health Sci 2017 Mar;6(1):17-24 [FREE Full text] [doi: 10.1016/j.jshs.2016.09.004] [Medline: 30356508]

21. Pasco D, Roure C, Kermarrec G, Pope Z, Gao Z. The effects of a bike active video game on players' physical activity and motivation. J Sport Health Sci 2017 Mar;6(1):25-32 [FREE Full text] [doi: 10.1016/j.jshs.2016.11.007] [Medline: 30356595]

22. Zeng N, Pope Z, Lee J, Gao Z. A systematic review of active video games on rehabilitative outcomes among older patients. Journal of Sport and Health Science 2017 Mar;6(1):33-43 [FREE Full text] [doi: 10.1016/j.jshs.2016.12.002]

23. Staiano A, Beyl R, Hsia D, Katzmarzyk P, Newton R. Twelve weeks of dance exergaming in overweight and obese adolescent girls: Transfer effects on physical activity, screen time, and self-efficacy. J Sport Health Sci 2017 Mar;6(1):4-10 [FREE Full text] [doi: 10.1016/j.jshs.2016.11.005] [Medline: 28491483]

24. Gao Z. Fight fire with fire? Promoting physical activity and health through active video games. J Sport Health Sci 2017 Mar;6(1):1-3 [FREE Full text] [doi: 10.1016/j.jshs.2016.11.009] [Medline: 30356540]

25. Ameryoun A, Sanaeinasab H, Saffari M, Koenig H. Impact of Game-Based Health Promotion Programs on Body Mass Index in Overweight/Obese Children and Adolescents: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Child Obes 2018;14(2):67-80 [FREE Full text] [doi: 10.1089/chi.2017.0250] [Medline: 29185787]

26. Benzing V, Chang Y, Schmidt M. Acute Physical Activity Enhances Executive Functions in Children with ADHD. Sci Rep 2018 Aug 17;8(1):12382 [FREE Full text] [doi: 10.1038/s41598-018-30067-8] [Medline: 30120283]

27. Budde H, Voelcker-Rehage C, Pietrabyk-Kendziorra S, Ribeiro P, Tidow G. Acute coordinative exercise improves attentional performance in adolescents. Neurosci Lett 2008 Aug 22;441(2):219-223 [FREE Full text] [doi: 10.1016/j.neulet.2008.06.024] [Medline: 18602754]

28. Hillman CH, Buck SM, Themanson JR, Pontifex MB, Castelli DM. Aerobic fitness and cognitive development: Event-related brain potential and task performance indices of executive control in preadolescent children. Dev Psychol 2009 Jan;45(1):114-129. [doi: 10.1037/a0014437] [Medline: 19209995]

29. Pesce C, Crova L, Cereatti R, Casella M, Bellucci M. Physical activity and mental performance in preadolescents: Effects of acute exercise on free-recall memory. Mental Health and Physical Activity 2009 Jun;2(1):16-22 [FREE Full text] [doi: 10.1016/j.mhpa.2009.02.001]

30. Wendt M. The effect of activity program designed with Intense physical exercise on the behavior of attention deficit hyperactivity disorder children. Dissertation Abstracts International Section A: Humanities & Social Science. 2000. URL: https://search.proquest.com/docview/304647212 [accessed 2020-09-15]

Abbreviations

EEG: electroencephalography
IMU: inertial measurement unit
ROM: range of motion
RPE: rated perceived exertion
SMR: sensorimotor rhythm
VR: virtual reality

XSL•FO
RenderX

Original Paper

Effects of a Personalized Fitness Recommender System Using Gamification and Continuous Player Modeling: System Design and Long-Term Validation Study

Zhao Zhao[1], PhD; Ali Arya[2], PhD; Rita Orji[3], PhD; Gerry Chan[2], MA

[1]Institute of Communication, Culture, Information and Technology, University of Toronto Mississauga, Mississauga, ON, Canada
[2]School of Information Technology, Carleton University, Ottawa, ON, Canada
[3]Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada

Corresponding Author:
Zhao Zhao, PhD
Institute of Communication, Culture, Information and Technology
University of Toronto Mississauga
3359 Mississauga Road
Mississauga, ON, L5L 1C6
Canada
Phone: 1 6137966008
Email: zhao.zhao@utoronto.ca

Abstract

Background: Gamification and persuasive games are effective tools to motivate behavior change, particularly to promote daily physical activities. On the one hand, studies have suggested that a *one-size-fits-all* approach does not work well for persuasive game design. On the other hand, player modeling and recommender systems are increasingly used for personalizing content. However, there are few existing studies on how to build comprehensive player models for personalizing gamified systems, recommending daily physical activities, or the long-term effectiveness of such gamified exercise-promoting systems.

Objective: This paper aims to introduce a gamified, 24/7 fitness assistant system that provides personalized recommendations and generates gamified content targeted at individual users to bridge the aforementioned gaps. This research aims to investigate how to design gamified physical activity interventions to achieve long-term engagement.

Methods: We proposed a comprehensive model for gamified fitness recommender systems that uses detailed and dynamic player modeling and wearable-based tracking to provide personalized game features and activity recommendations. Data were collected from 40 participants (23 men and 17 women) who participated in a long-term investigation on the effectiveness of our recommender system that gradually establishes and updates an individual player model (for each unique user) over a period of 60 days.

Results: Our results showed the feasibility and effectiveness of the proposed system, particularly for generating personalized exercise recommendations using player modeling. There was a statistically significant difference among the 3 groups (full, personalized, and gamified) for overall motivation ($F_{3,36}$=22.49; P<.001), satisfaction ($F_{3,36}$=22.12; P<.001), and preference ($F_{3,36}$=15.0; P<.001), suggesting that both gamification and personalization have positive effects on the levels of motivation, satisfaction, and preference. Furthermore, qualitative results revealed that a customized storyline was the most requested feature, followed by a multiplayer mode, more quality recommendations, a feature for setting and tracking fitness goals, and more location-based features.

Conclusions: On the basis of these results and drawing from the gamer modeling literature, we conclude that personalizing recommendations using player modeling and gamification can improve participants' engagement and motivation toward fitness activities over time.

KEYWORDS

XSL•FO
RenderX

Introduction

A sedentary lifestyle is defined as a lifestyle in which an individual does not receive regular amounts of physical activity, which is becoming a significant public health issue [1]. Various solutions have been considered to encourage a more active lifestyle. Among them, combining exercise with gameplay [2] and the use of wearable trackers to motivate and recommend physical activities [3] have received widespread popularity, but both have user retention issues [4]. This paper addresses the issue of improving long-term engagement with such physical activity recommenders and exercise games.

The popularity of computer games and their engaging nature has created a strong trend to use games for nonentertainment purposes [5]. This trend includes overlapping topics and terms such as gamification (the use of game features and mechanics in nongame applications [6]), serious games (aimed primarily at being an educational yet entertaining tool [7]), persuasive games (games for promoting behavior change [8]), and exergames (a combination of physical exercise with games [9]). In particular, gamification has received significant attention because it can be seen as an umbrella topic covering a range of options from implementing few game elements (such as leaderboards) in regular activities to performing serious tasks as a full game [10,11].

Games and gamified activities are effective persuasive tools for motivating human behavior [12]. Although recent years have seen an increase in persuasive applications designed to promote more active lifestyles [13], studies have suggested that a *one-size-fits-all* approach is ineffective for such persuasive applications because different users are motivated by different persuasive strategies [14] as well as personal reasons such as curiosity and social rewards [15]. There is an increasing demand for personalization as a means of tailoring an experience to individual needs and interests [15,16]. This is particularly the case for persuasive and recommender systems such as those in marketing, education, and health, where retention is as important as initial action [17]. Among different personalization solutions, player modeling in games that aims to understand players to enhance game experience has been an active research topic [18]. It aims to describe a game player's traits and preferences as well as the players' cognitive, affective, and behavioral patterns [19] within well-defined structures that allow designers to tailor game contents or goals automatically to suit the needs or preferences of individual players.

Gamification has rapidly emerged over the past years, especially in the area of exercise and fitness [20], as a tool to promote healthy behaviors and maintain an active lifestyle [9]. Researchers have used various gameplays and game features to make exercise and physical activities more engaging and attractive [2,3]. The use of gamification to promote a more active lifestyle can be through either formal exercises performed as games (ie, exergames) or combining games with other physical activities that are not as rigorous as formal exercise (ie, a walk to work). In this paper, we refer to all these cases as gamified physical or fitness activity and use the term "exergame" loosely to indicate the same type of activity. These activities have the potential to help users achieve their fitness goals and increase engagement and pleasure by adding game features to physical activity [12,19].

However, most existing work on gamification and persuasive games in health and wellness are limited because of their use of one-size-fits-all approaches, which have been shown to be suboptimal [21]. As discussed in the next section, there are some research efforts that have used more in-depth personalization, but they are not focused on exergames, and initial attempts at personalization are mostly limited to a small set of static demographic parameters about the user (such as age, gender, and occupation), which makes them more *categorized* rather than *personalized*. In addition, their effectiveness in promoting the desired behavior was mostly evaluated based on a single point of use and feedback (short term). Few attempts have been made to resolve these issues. For example, MyBehavior [22] used a tracking-based and dynamically modified user model to recommend activities, but the system is not gamified and is based on a limited model of daily activity information. A proper combination of a detailed model with features such as personality types, modeling-based personalization, and recommendation with adaptive gamified elements is still missing in the area of exergames, as discussed in more detail in the next section.

To address these research gaps, we propose a comprehensive model for gamified fitness recommender systems that use detailed and dynamic player modeling and wearable-based tracking to provide personalized game features and activity recommendations. We also present the results of a long-term investigation on the effectiveness of our recommender system that gradually establishes and updates an individual player model (for each unique user) over a relatively long period (60 days).

We hypothesized that (1) player modeling based on continuous player activity tracking is an effective approach for personalizing activity recommendations to individual players and (2) combining player modeling and gamification can promote long-term engagement with the system.

On the basis of these hypotheses, we aim to address the following specific research questions in this paper:

1. How can we generate and use continuous player modeling to personalize activity recommendations for each user?
2. Can the combination of player modeling and gamification techniques improve user engagement and experience toward fitness activities over time?

To achieve this, we designed a new player model and a related system architecture for a gamified fitness activity recommender system. To evaluate the effectiveness of the model-driven gamified fitness activity recommender system, we conducted a long-term study on 40 participants and examined the effectiveness of our gamification approach in promoting physical activity in comparison with a control group. We randomly assigned our participants into 4 distinct groups corresponding to 4 experimental conditions:

1. The full group received the application with both gamified features and personalized recommendations (based on player modeling).

2. The personalized group received only personalized recommendations but no gamified features.
3. The gamified group received nonpersonalized recommendations with gamified features added.
4. The control group received generic nonpersonalized recommendations and nongamified features.

The results of a 60-day-long study showed that the idea of generating personalized exercise recommendations using player modeling is feasible. Moreover, it showed that personalizing recommendations using player modeling in combination with gamification can improve participants' level of engagement and motivation toward fitness activities over time.

Our work includes the following major contributions to the field of fitness recommenders and exergames:

1. We offer a conceptual model and system architecture for personalized gamified activity recommendations using a combination of player modeling, gamifications, and activity tracking.
2. We build a comprehensive player model for personalization that is dynamically updated by continuously tracking player contexts.
3. We design and validate a 24/7 recommender system for personalized activities combined with various gamified elements that adjust to a player's and environmental contexts.
4. Finally, to demonstrate the feasibility of our approach, we conducted a long-term (60-day) field study with 40 participants to evaluate the proposed system and compare the effectiveness of the 4 experimental groups.

Through the design and development of a new recommender system and a long-term study, we hypothesized and evaluated the effects of gamification and player modeling on physical activity. To the best of our knowledge, this research is the first to link research on player modeling, gamification, and activity recommendation to propose an approach for personalized activity recommendation that is continuously and dynamically updated to reflect users' changing contexts and states.

Related Work

Gamification

The motivational effects of gamification have been widely studied by researchers. It has been shown that common game elements such as badges, rewards, leaderboards, and avatars are commonly and successfully used to motivate players [10,21-23]. On the other hand, researchers have argued for various areas of improvement in current gamification research and applications, including diversity in themes and context, study duration, and sample size [6,23], increasing motivation by relying on more intrinsic factors [22,24], continuous adaptation [25], and personalization [6]. In the study by Loria [25], a framework for improving the player experience and customizing content generation is proposed by continuously monitoring how players interact during the game by analyzing information such as players' in-game behavior and players' social network. Other researchers have also tackled adaptive gamification. For example, Böckle et al [26] identified 4 main elements as the basis for defining meta-requirements and designing principles

for building an adaptive gamification system: (1) consider the purpose of adaptivity, (2) define the adaptivity criteria, (3) design the adaptive gamification mechanics and dynamics, and (4) design meaningful adaptive interventions. The researchers applied this framework to the design of a web-based platform for knowledge exchange in postgraduate medical training and reported positive user acceptance, feedback, and increased usage.

Nicholson [24,27] discussed the idea of meaningful gamification (or playification), which focuses on playfulness and activities that make sense to each player and rely on intrinsic motivations as opposed to following specific rules to win. Bertran et al [28] built on this idea and proposed the situational play design, which is a framework for designing context-based and personalized games. Orji et al [14-16], among others, as discussed in a later section, also discussed the idea of player modeling for designing more effective serious and health games. Overall, research in the gamification domain suggests the need for more personalized games that depend on the players' context and their motivations. The research presented in this paper expands on these ideas and addresses some of the identified needs, such as tracking and understanding the player using a comprehensive individual-level model, personalizing both game features and recommended activities in exergames, and performing long-term studies, within the context of gamified physical and fitness activity recommenders. We discuss the research in these specific areas in later sections.

Gamified Physical Activities

Despite their positive effect on promoting an active lifestyle, gamified physical activities face the problem of sustainability (also referred to as player retention here and in other literature) [20,25,26]. Although players may feel excited and motivated to play at first, over time and sometimes quickly, they may lose their willingness to continue. There are studies focusing on the motivation and sustainability of exergames and gamified fitness activities. For example, Campbell et al [21] discussed the concept of everyday fitness games and suggested that for applications that people frequently use in their everyday lives, the design needs to be fun and sustainable as well as adapt to behavioral changes. Macvean et al [29] reported a 7-week study on users' physical activity, motivation, and behavioral patterns using exergames and suggested that longitudinal studies are necessary for evaluating motivational effects as exergames ensure that the intensity of a user's behavior is appropriate and sustained. Previous work [30] also showed that based on existing technologies and user needs, the idea of employing wearable activity trackers for gamification of exercise and fitness is feasible, motivating, and engaging. Adding dynamic features could have a positive impact on user motivation toward the gamified exercise system, and the gradual release of application features could increase the user retention rate. It was also found that each user was unique and motivated by different types of game features. Therefore, based on these results, it seems reasonable to generate customized workout sessions to fit different user fitness conditions and interests.

Recently, significant research has been devoted to the design of active games (also commonly referred to as exertion games or exergames) to match the needs of specific groups such as

those with disability [28,29] or senior citizens [31] or to use specific technologies such as virtual reality (VR) [32]. Particularly in the context of VR-based exergames for adolescents, research shows that game elements such as the use of rewards, increasing challenge levels, frequent updates, and social or multiplayer options are important aspects for continued engagement in physical activity [33]. Researchers have also investigated design principles for active games [34,35]. Although impressive and invaluable in their findings, these efforts primarily focus on the design of particular games, game features, and actual gameplay mechanisms that are best suited for increasing physical activity for the target group or individual. On the other hand, in 24/7 activity recommenders, the focus is less on designing a particular game and more on gamifying the daily experience and recommending activities based on the daily routine using dynamic player modeling.

Player Models

Busch et al [30] indicated that the *one-size-fits-all* approach does not work for persuasive game design. Thus, player-type models could be used when tailoring personalized persuasive systems. One of the most frequently used player-type models is the one developed by Richard [36], who identified 4 player types and proposed that each player has some particular preference for one of the types, which makes them mutually exclusive. Another model is the BrainHex model [37], which is a relatively new model but has been validated using a large pool of participants [38]. In BrainHex, player types were not mutually exclusive. Scores under each category are presented to determine the player's primary type and subtypes. It also connects player types to the game elements. Moreover, the Hexad model [39], which is of particular interest in our work, is a gamification player-type model created for mapping user personality onto gamified design elements. We considered using the Hexad model in our player model because it specifically targeted gamified systems. It proposes 6 player types, and the player types of individuals are correlated with their preferences for different game design elements. Design guidelines for tailoring persuasive gamified systems to each gamification player type have also been studied [17].

Furthermore, Wiemeyer et al [40] discussed the concept of player experience (with a focus on individual) versus game usability (with a focus on technology) and reviewed various theoretical models that can help understand the player experience. These models are particularly helpful when designing full games as opposed to gamifying everyday activities, which is the goal of this research. However, their insights, such as an integrative multidisciplinary model of player experience, can be helpful in future phases of our research when we focus on the design of game elements. For the work presented here, our focus was primarily on showing how the combination of gamification and player modeling could improve engagement. Better player experiences can be achieved through more complicated models and game features that are beyond the scope of this work.

Personality type also plays an important role in determining people's fitness tastes [41]. Some people may prefer swimming laps solo, whereas others may enjoy attending a rowdy group-cycling class. These preferences have less to do with people's physical characteristics and are affected more by personalities. Matching activities to personality type has been shown to have real-world relevance [42]. Research suggests that people who engage in personality-appropriate activities will stick with the activities longer, enjoy their workout more, and have a better overall fitness experience [43]. Brue [44] created a system based on the principles of the Myers-Briggs–Type Indicator (MBTI) assessment. She used MBTIs and reworked them into an easily maneuverable color-coded fitness personality model, the 8 Colors of Fitness, which is also used in our player model. Each color is associated with 2 personality types from the 16 possible MBTI types [45]. For example, blues are loyal, traditional, dependable, and straightforward, whereas greens are nature lovers who seek to quietly merge with the outdoors [42].

Moreover, in recent years, the use of wearable sensors in human activity recognition has become popular [46], in which most of the measured attributes are related to the user's movement (eg, using accelerometers or GPS), environmental variables (eg, temperature and humidity), or physiological signals (eg, heart rate or electrocardiogram). These data types are naturally indexed over the time dimension, consistent, and convenient to access, which could be used in modeling and predicting a user's daily activity pattern.

Although there is a significant amount of research on the subject of player modeling, none of the existing studies have examined how to use a comprehensive player model. In addition, no previous research has simultaneously considered both game features and recommended activities in exergames design and investigated whether it is an effective approach over the long term.

Personalized Activity Recommendations

Personalized recommender systems for physical activity have been studied by many researchers. For example, Guo et al [47] proposed a system that recognizes different types of exercises and interprets fitness data (eg, motion strength and speed) to an easy-to-understand exercise review score, which aims to provide a workout performance evaluation and recommendation. Although it achieved 90% accuracy for workout analysis, it focuses only on recognizing fitness activities and not personalizing or gamifying them. He et al [48] introduced a system designed to be context aware for physical activity recommendations. It focuses on selecting suitable exercises for individualized recommendations. A smartphone app was developed that could generate individualized physical activity recommendations based on the system's database of physical activity. The focus of their work is to recommend different types of activities but does not take into account personal details such as proper time, location, and intensity or any gamified elements.

Broekhuizen et al [49] proposed a framework called PRO-fit, which is another example that employs machine learning and recommendation algorithms to track and identify users' activities by collecting accelerometer data, synchronizes with the user's calendar, and recommends personalized workout sessions based on the user's and similar users' past activities, their preferences, and their physical state and availability. The authors highlighted

that many applications nowadays are more focused on tracking user activities but do not provide a recommendation system that would help users choose from activities based on their interests and accomplishment of goals. Therefore, the authors were motivated to design a personalized fitness assistant framework that acts as a motivator and organizer for fitness activities, making it easier for users to create and follow their workout plan and schedule the sessions according to their availability and preference. Compared with PRO-fit, our proposed system provides recommendations in real time throughout daily life as opposed to the prefixed recommendations that are not based on any player or exerciser-type model employed in PRO-fit.

Mittal and Sinha [50] used personal information to recommend general activities such as visiting attractions and shopping. Although not focused on fitness activities or gamification, their notion of modeling user data as the base for recommendation is in line with our proposal. Ni et al [51] used a variety of user data such as daily routine and heart rate to recommend workout routes. Their method is more focused on physical activity recommendations but is limited to recommending routes and does not include gamification elements. Similarly, Rabbi et al [22] proposed MyBehavior, which is a system for tracking users using mobile devices and suggesting food and physical activities. MyBehavior provides personalized and real-time suggestions but is not gamified and does not include an explicit player model. As such, it does not take advantage of full personalization or more engaging features that a game can offer. In line with this, Ghanvatkar et al [52] conducted a comprehensive review of user models used in recommender systems. They highlight that activity profile, demographic information, and contextual data such as location are among the top items to include in user models. In this research, we have defined our player model to include gamer information using Hexad and demographic, activity, and exercise submodels, as suggested by Ghanvatkar et al [52].

Summary of Research Gaps

As reviewed in previous sections, research on gamified physical activities and related topics has achieved significant results but requires more work to fill the existing gaps. We identified that the main research gap within the context of exergames is the notion of *personalized gamification* (a combination of gamified physical activities with player model–based personalization), including understanding the players and their environment and adapting the game features and physical activities dynamically. None of the existing studies have successfully investigated the effect of gamification and personalization individually with respect to promoting the efficacy of an intervention, specifically a physical activity intervention within a single application. In addition, long-term studies outside controlled environments and

real-time activity tracking and recommendation are also frequently missing in existing research on personalization and exergames.

However, the existing body of research provides invaluable insight into recommendations based on real-time tracking, important parameters to include in a player model, and the design of exergames in general. Expanding on existing studies and trying to fill the abovementioned gaps, we propose a conceptual model and system architecture that bring together game elements, dynamic player modeling, and activity tracking to personalize exergames in terms of both game features and recommended activities. We built a comprehensive and dynamic player model for personalization that is continuously updated by tracking the player and offers 24/7 personalized activity recommendations. Finally, we conducted a long-term user study in the wild to evaluate the proposed system.

To the best of our knowledge, although some of the features of our study have been suggested and/or investigated by others, no long-term comprehensive study has been conducted to integrate and evaluate them in real-life exergame apps.

Methods

Conceptual Model

Although the existing studies have addressed many aspects of these diverse fields, as discussed earlier, they have not been properly integrated to develop engaging and sustainable exergames. For example, the effect of various game features and continuously adapting the game to player needs and interests have not been investigated in the context of exergames.

In this section, we describe our proposed system architecture and related research methods. This proposal is based on our new conceptual model developed after reviewing related work, consisting of the following principles:

1. Advances in wearable technologies allow game designers to use commercially available activity tracking sensors and mobile devices as a major element of exergames.
2. A game with a static design, no matter how interesting, will lose its attraction after a while. As such, it is important to *add new features* over time to keep players engaged.
3. Although different methods exist for adding dynamic features to games, designers have a limited ability to provide new features constantly, and there is no guarantee that they will be attractive to users. An alternative (or complementary) approach is to dynamically modify the game by *adapting to the player*.

Our conceptual model, which builds on our previous work [28] along with 2 new components, is illustrated in Figure 1.

Figure 1. Conceptual model of the proposed system. Components surrounded by parallelograms were explored in our previous work, whereas components surrounded by rectangles are the aim of this research for personalizing physical activity recommendations. Arrows show the connection between the components.
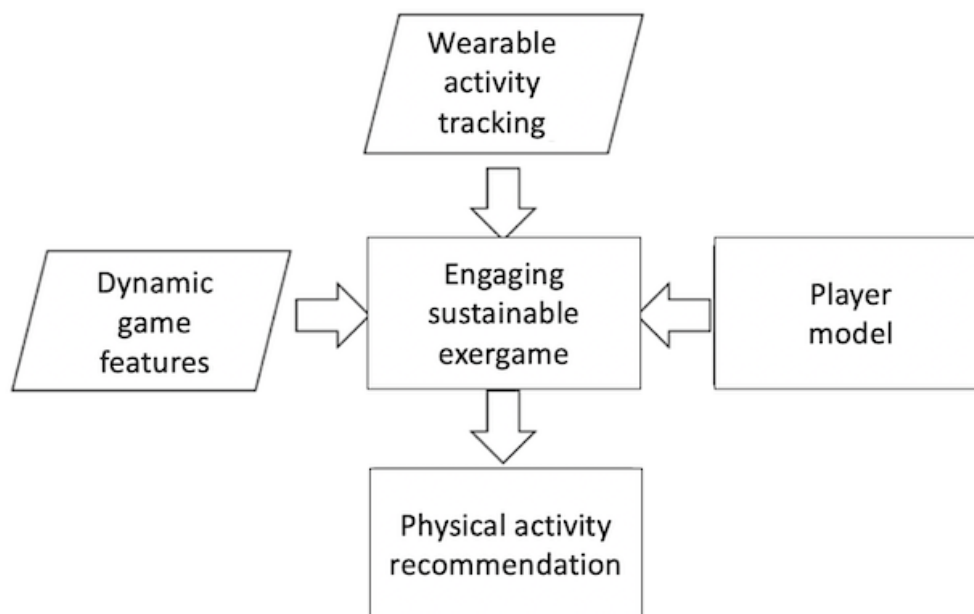
In our previous work [53], we investigated the effect of using wearable activity tracking in exergames and the long-term effectiveness of using a dynamic game feature–releasing system in sustaining exergames (marked parallelograms). In this paper, we aim to further investigate the gamified features for increasing retention, exploring 2 additional components: (1) player modeling in the personalization of exergames and (2) how to use such a system to generate personalized physical activity recommendations (marked rectangles). The arrows demonstrate how each of the components are related and can directly influence each other. We believe that tracking activity using wearable technology, providing dynamic game features, and detecting a player's preferences using a player modeling approach can all contribute to creating a more engaging exergame experience, which in turn can generate more personalized physical activity recommendations that players will likely find motivating and satisfying.

System Design

On the basis of our proposed conceptual model, bringing wearable activity trackers or smartwatches into exergames, dynamically updating game features, and using player modeling for personalization of exergames is being proposed as a solution to the research problem. Therefore, a wearable-based exergame with a comprehensive player model for personalization, recommending customized activities, is proposed as a potential system for further investigation.

The proposed system contains 3 main components: a player model, a recommendation engine, and a game generator. The player model takes different types of user data and predicts user preference for physical activities and finds the proper time and location for recommending activity sessions. It consists of several submodels that cover the user's general, personality, and daily activity data. The recommendation engine used the output of the player model and generated customized physical activity session recommendations for individual users (including the proper time, location, intensity, and potential type of physical activity). The game generator adds customized game elements to the recommendation and generates the final game content that users can interact with. Wearable activity trackers or smartwatches are used in the system to track the user's activity and introduce diverse interactions. The combined use of mobile apps and wearable apps will allow users to interact with the system with different modes. The detailed design and development of the system are introduced in the following section.

Overall, a wearable-based exergame system, with a comprehensive player model for physical activity recommendation and game customization, is proposed as a solution to the exergame retention problem.

Application Design and Implementation

On the basis of the proposed system architecture, a Wear OS (formerly Android Wear) application is implemented as the user interface (UI), which tracks the user's activities and provides gamified fitness recommendations. The overall architecture of the application is illustrated in Figure 2.

Figure 2. App architecture. How data were collected and transferred to each submodel of the system and how they were used to generate recommendations and game content. Info: information.
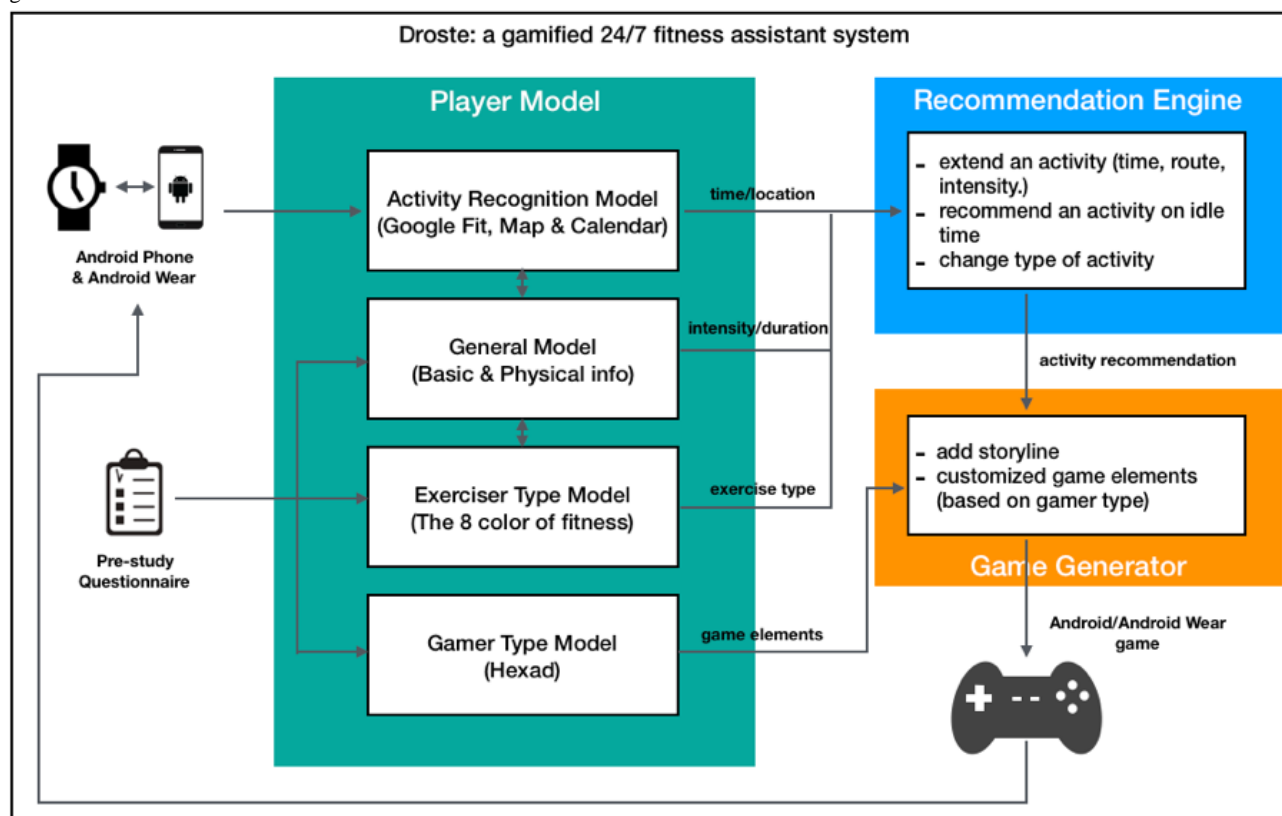
In this application, and based on the conceptual model in Figure 1, the player model consists of 4 submodels: (1) an activity recognition model that tracks player activities, (2) a general model that holds basic information about the player, (3) an exerciser-type model that includes information required for recommending activities, and (4) a gamer-type model that is used to choose game features. Each of the submodels is mainly

in charge of generating one part of the recommendation, as shown in Table 1. Choosing game features and physical activities are the 2 main personalization options and each has its own submodel. Tracking daily activities is an essential part of the system, which also has a submodel. The fourth submodel holds general player information, such as gender, age, weight, and height.

Table 1. The roles of each submodel.

Submodel	Role
Activity recognition model	Time and location
General model	Intensity and duration
Exerciser-type model	Exercise type
Gamer-type model	Game elements

Although each submodel is designed to generate one particular part of the recommendation, they are still connected to each other to create a more reliable overall recommendation. For example, the exerciser-type model is built for each individual user for recommending different types of activities based on their personalities but it also relies on the general model, which is built based on a user's fitness and health condition, to exclude those activities that may be suitable for their personality type but not for their health condition. We refer to the theoretical foundations from the Global Recommendations on Physical Activity for Health (GRPAH) [54] to determine proper exercise recommendations in nonpersonalized cases. The GRPAH is an accepted tool approved by the World Health Organization for recommending the appropriate exercise type, duration, and

intensity. We use the 8 Colors of Fitness model [44] to suggest different types of activities for the personalized groups. This model is one of the few that uses personality type as the basis of activity recommendations and is suggested by other researchers and practitioners [49,50].

The recommendation engine is a decision tree–based module that uses all the information generated from the player model to create personalized recommendations for each individual user. It could either extend an existing activity (eg, by recommending a longer exercise time, a longer running path, or appropriate intensity), recommend some activities on the user's idle time, or simply recommend a different type of activity. An example of a decision tree is illustrated in Figure 3. As the recommendation system for physical activity itself is

a relatively complicated topic, we do not consider it as the highest priority for this study. Therefore, we only employed simple decision tree methods to generate basic recommendations (we have ensured that all the recommendations followed the GRPAH guideline for daily physical activities). We are aware that the rationality and quality of the recommended activities would have an impact on user experience. Therefore, after verifying the feasibility of the proposed idea and the roles personalization and gamification performed in this type of system, our research goal will be to investigate the recommender system of physical activities.

Figure 3. Example of the decision tree used in the recommendation engine (simple version). PA: physical activity.
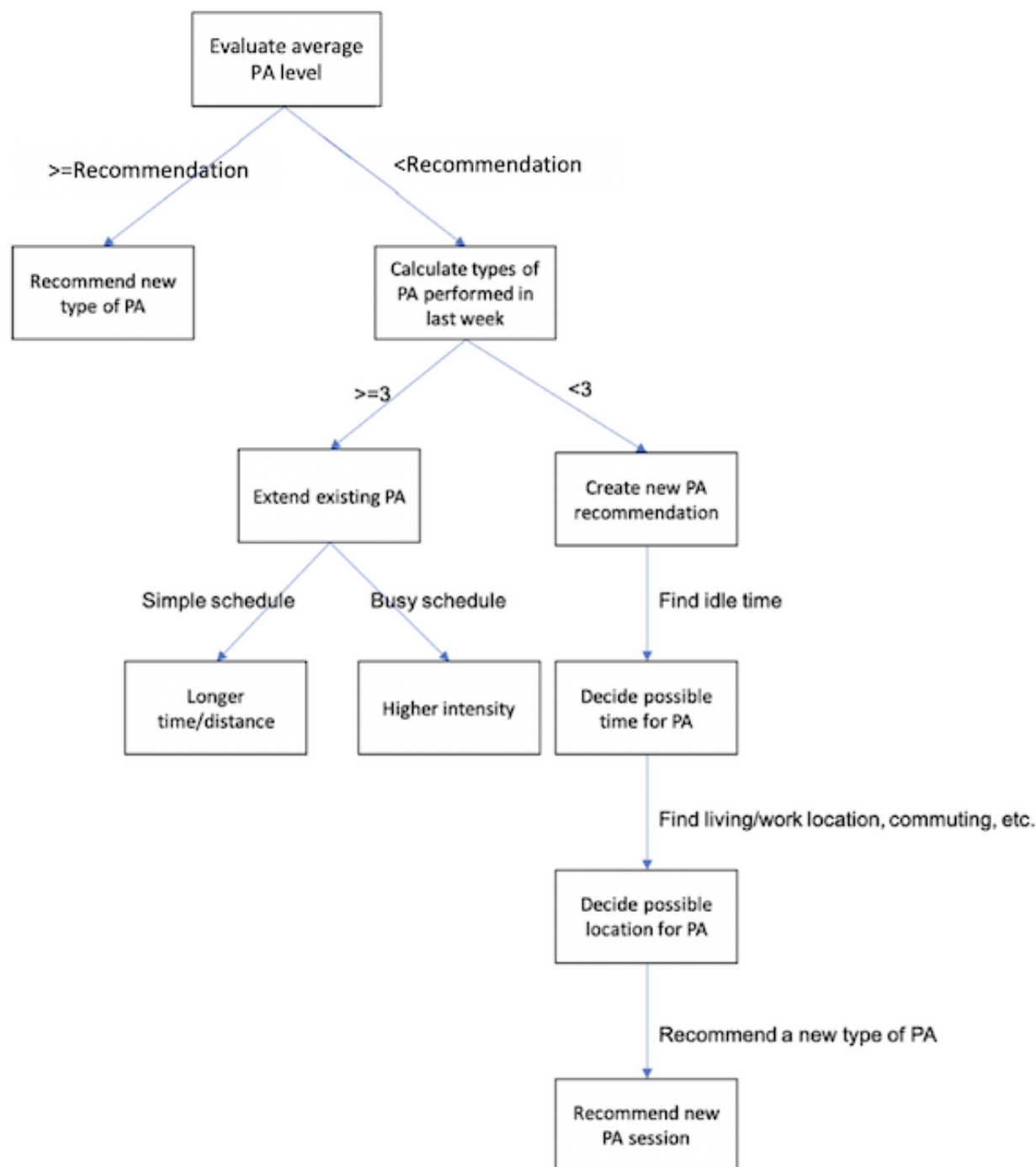

The game generator is responsible for adding game elements to the recommendations to gamify the activity suggestions generated by the recommendation engine. The type of game elements to be added is determined by the Hexad player–type model [39]. Our work is also partly based on Orji et al [17], as it adopted a similar Hexad player model. However, as opposed to using the persuasive strategies recommended by Orji et al

[17], we used the game elements recommended by the Hexad player–type model, which is more in line with our objective in this work, designing gamified physical activity recommendations. Details of the game and activity recommendations are provided later in this section.

A Wear OS app was developed for this study. The app is a conversation-based game in which all the interactions happen in the form of a conversation between the user and the *future self*. The game is based on a story in which a 1-day user receives a message from the future self in 20 years telling him or her that the world is about to end in that future world but only the user can save it by completing a series of tasks. Then, the future self will guide the user through daily activities, which are generated by the recommendation system in a gamified structure. The choice of this game was informed by our need to have a simple design that is capable of incorporating our research requirements but at the same time is not too complicated to

develop with many possible confounding factors. We also did not want to introduce various esthetic and design variables to the study that may interfere with our studied research variables and influence our results. For the same reason, we did not try to incorporate our study within an existing game, even though adding these features to games that the user prefers may be another motivating factor in the future. However, it is essential to establish their effectiveness first in isolation. The current system UI was created using a rapid prototyping approach. A pilot study was also conducted before the formal study to ensure that the labels and buttons are clear. The main UI and app icons are shown in Figure 4.

Figure 4. Example app interface and icon. (a) A snippet of one conversation between the system and the user. (b) A display of new mission for the user that he or she can choose to accept or decline. UI: user interface.
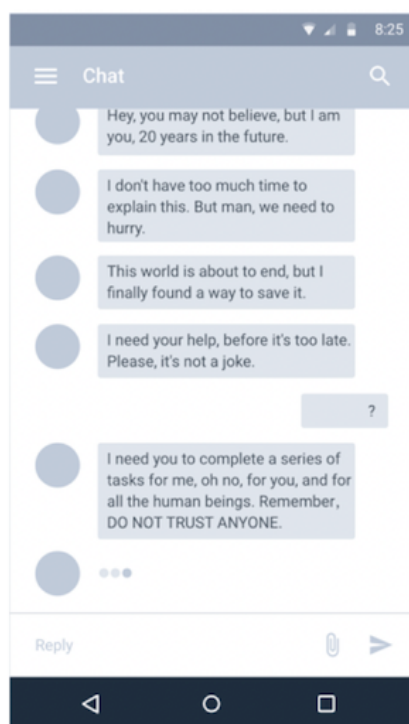
(a) Android UI
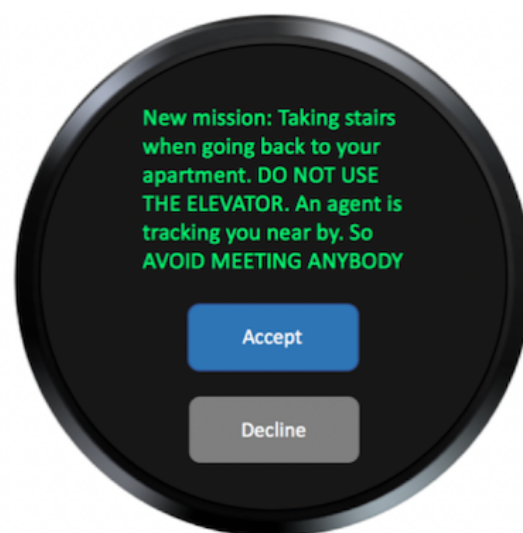
(b) Wear OS UI

The app tracks the user's daily activity through Android ActivityRecognition [55] and Google Fit Application Programming Interface (API) [56], which allows up to 6 user activities to be recognized in real time: in vehicle, on foot, running, walking, on bicycle, and still.

The Google Fit API provides encapsulated daily activity–related data such as calories burned, daily steps, and heart rate history (if applicable) tracked by both phone and watch sensors. All the collected activity data, along with their time stamps and location information, are used as input features to train a daily activity model for each individual user by which possible exercise time and location are predicted. As shown in Figure 4, the app is a conversation-based game. We used Wit.ai [57] to generate storylines and to build a bot that can talk to participants and perform some general greetings, tell the time, and talk about the weather. Wit.ai is a tool that uses natural language

processing to understand human language and we used its message API to create a chatbot, which aims to understand a user's intents and lead participants to designed storylines. Moreover, we have included a weather assistant in the system (through the Weather API [58]) to help participants in planning activities around the weather.

When designing the game features, we employed the Hexad player types [39] and the game design elements guide [17]. Hexad suggests that game design elements are preferred by each player type and we implemented 1 element for each type of user in this study for a personalized game experience (in addition to the game storyline). We integrated the following gamification elements in our game (Table 2). Figure 5 shows some screenshots of example game elements for different player types. If there was a tie in scores between the 6 types, we randomly chose 1 element of the highest score to add.

Figure 5. Example game elements: (a) profile in daily view (including points and challenges), (b) profile in weekly view (including points and challenges), (c) connect to Facebook view, (d) hacker mode view, and (e) theme color customization view.

Table 2. The motivation and corresponding game elements added for each type of player.

Hexad type	Motivation	Game element
Socializers	Relatedness	Link to social network
Free spirits	Autonomy and self-expression	Theme color customization
Achievers	Mastery	Challenge
Philanthropists	Purpose and meaning	Game experience sharing
Players	Rewards	Points
Disruptors	Change	Hacking mode

Link to Social Networks

Socializers are motivated by relatedness. They want to interact with others and create social connections [39]. Therefore, we provided them with an interface for linking the game to their social network as their unique feature so that they could share their game performance or achievements to their Facebook page, team up with those friends who are already in the game, or invite new players to the game.

Theme Color Customization

Free spirits are motivated by autonomy and self-expression. They want to create and explore the game and prefer features such as unlockable content and customization [17,39]. Thus, we added a feature of theme color customization so that they could customize their game UI by unlocking different themes.

Challenge

Achievers are motivated by mastery. They are looking to learn new things and want to overcome challenges [39]. Therefore, we added a challenge system for them in our game, in which tasks were assigned to them as challenges.

Game Experience Sharing

Philanthropists are motivated by purpose and meaning. They want to give to other people and enrich the lives of others in some way with no expectation of reward [39]. For philanthropists, we added a feature for them to share their game experience with other players. A forum-like interface was added to their version of the game in the main screen that allowed them to browse and answer questions of other players. They also receive notifications when there is a new question in the forum.

Points

Points have been shown to positively affect *players* [17,39]. They will do what is needed for them to collect rewards from a system. For players, points in our game can be collected and used as virtual currency to buy extra themes or virtual equipment.

Hacking Mode

Disruptors are motivated by change. In general, they want to disrupt the system [39]. We added a hacking mode for disruptors, in which they can use the command-line interface to access their own game database to make changes to the storyline or delete their records of the game and, eventually, they may *destroy* the system.

As mentioned, the application is a Wear OS game that requires combining the use of both an Android phone and an Android watch to optimize its recognition accuracy and gamified experience. For activity recognition, our app uses watch sensors for better accuracy. However, in many situations, participants may choose not to wear the watch. In those cases, when the watch was not connected, we used the built-in phone sensors instead such that the game could run individually on the phone without the watch. A phone clearly offers more screen space and abilities, such as typing messages, compared with a watch.

User Study Design

Multiphase Research

The proposed conceptual model and the system were evaluated based on a multiphase user study. In our previous work [53], we introduced phases 1 and 2, which can be summarized as follows:

For research phase 1, in-lab user tests of 20 participants were conducted to evaluate the effectiveness of the combined use of games and wearable devices in promoting exercise and to investigate the usability of the proposed approach and the effects of different factors within the system.

In research phase 2, a 70-day user study of 36 participants was designed to verify the hypothesis that adding different game features and gradually releasing them can positively affect user engagement and retention.

In this paper, we present research phase 3, which is a 60-day long-term study with 40 participants, to demonstrate the feasibility and effectiveness of using a player modeling technique in the personalization of exergames.

Participants and Groups

A total of 40 participants were recruited locally from the Ottawa area by posters as well as via the web through the Android Wear Forum [59]. Of the 40 participants, 23 were men and 17 were women. Their average age was 26.93 years, with an SD of 6.07 years. We randomly divided our participants into 4 groups based on the versions of the app they received: full (gamified and personalized), gamified only, personalized only, and the control (neither personalization nor gamified, as the control group). Participants were randomly allocated to groups and the distribution with respect to exercise or player type and physical activity level did not seem to be particularly biased (Table 3). Participants' physical activity levels were collected at baseline before beginning the study.

Table 3. Demographical data for the 4 participant groups (N=40).

Characteristic	Participant group			
	Full	Gamified	Personalized	Control
Age (years), mean (SD)	24.93 (7.27)	26.65 (5.58)	27.85 (6.26)	25.78 (5.93)
Gender, n (%)				
Male	6 (60)	6 (60)	5 (50)	6 (60)
Female	4 (40)	4 (40)	5 (50)	4 (40)
Hexad user types, n (%)				
Philanthropist	1 (13)	1 (11)	1 (10)	1 (11)
Socializer	1 (13)	2 (22)	1 (10)	2 (22)
Free spirit	1 (13)	3 (33)	2 (20)	2 (22)
Achiever	2 (25)	2 (22)	3 (30)	2 (22)
Disruptor	1 (13)	0 (0)	0 (0)	1 (11)
Player	2 (25)	1 (11)	3 (30)	1 (11)
8-color personalities, n (%)				
Blue	3 (30)	2 (20)	1 (10)	2 (20)
Gold	1 (10)	1 (10)	1 (10)	2 (20)
White	2 (20)	1 (10)	3 (30)	0 (0)
Purple	0 (0)	1 (10)	1 (10)	0 (0)
Green	1 (10)	3 (30)	1 (10)	2 (20)
Red	1 (10)	0 (0)	1 (10)	2 (20)
Saffron	0 (0)	0 (0)	2 (20)	1 (10)
Silver	2 (20)	2 (20)	0 (0)	1 (10)
Physical activity level (hours per week), mean (SD)[a]	4.04 (2.35)	3.95 (3.21)	4.82 (2.53)	3.83 (2.92)

[a]Physical activity levels were self-reported at baseline.

To increase the duration in each group and reduce the chance of groups affecting each other, all participants remained in the same group for the entire study duration rather than randomly trying all 4 groups.

The recommendations for the control group and the gamified group were created based on established exercise guidelines and were reasonable recommendations for the general population. To ensure this, we referred to the theoretical foundations from the GRPAH [54] to determine proper exercise recommendations. Our choice for nonpersonalized groups closely follows the one-size-fits-all recommendation method, which has been generally used in most physical activity recommendation applications, such as Apple Watch (recommends a 30-min walk per day) or Fitbit (daily 10,000 steps).

Furthermore, the main purpose of our study was to demonstrate the effectiveness of personalized recommendations. Although we tried to offer a reasonable experience for nonpersonalized groups, the effectiveness of personalization, especially in terms of recommendation, was our hypothesis when other variables are held constant. Therefore, we did not use existing commercial apps for comparison in this study because we tried to avoid bringing in possible extraneous or confounding variables such as esthetic and gameplay features that were not our focus.

Figure 6 shows an example of how recommending the same 30-min walking activity will look for the 4 study groups. The full group received the recommendation through a gamified story (guided by the future self) with the game element of challenge based on their player type of achiever and a personalized walking path. The personalized group also received a personalized route but no game story or elements. The gamified group received no personalized route but had the game story and the game element of points (randomly assigned because no player model was used for the gamified group). The control group received no personalization or gamification as a control group. In the screenshot of the control group, we showed an example of how the weather assistant worked. Note that the example conversations were from screenshots and some details related to the context were not fully displayed.

Figure 6. Example of recommendations for different groups: (a) full, (b) personalized, (c) gamified, and (d) control.

To control the gamification level between groups, for the gamified group, as there was no player model used, we randomly assigned a game element from Table 2 to each participant to bring them to the same gamification level as the full group. As the members of personalized groups received different game elements based on their individual player model, we decided that a random selection for nonpersonalized groups would be the closest nonpersonalized option.

We also limited the number of personalized recommendations to 2 times a day to eliminate the variability of engagement caused by frequent recommendations. The gamified and control groups (without personalization) received 2 messages per day at 9 AM and 5 PM. We chose these 2 times because 9 AM is the time of day that most of our participants were active. We did not send the notification earlier because we did not want their sleep to be interrupted. We chose 5 PM because most people are off from work at 5 PM. The full and the personalized

group received messages based on when they got up and when they left work, as recorded in their individual player model. The results presented in this paper are based on data from a 60-day experiment.

Procedures and Data Collection

The study was approved by the research ethics board. We asked participants to complete a prestudy questionnaire before providing them the app (Multimedia Appendix 1). The questionnaire asked demographic questions including age, gender, height, weight, number of hours they spend per week exercising, type of Android Wear owned, and types and duration of playing video games (eg, PC, console, and mobile). Two web-based questionnaires were provided and participants completed them, which provided us with the results to determine their player and exerciser type [60,61]. The app was distributed to participants through the HockeyApp (now Microsoft Visual Studio App Center) [62] after receiving participants' gamer and exerciser-type results. Application features were selected based on the participant's player model.

For in-game data collection, we used Google Analytics API [63] to track all participants' comprehensive in-app behavior data, including screen views and tapped events with associated timestamps. We used Google Fit API to track user daily activity data and a pop-up question asking participants if the recommendation they received that day was useful. The notification was sent to participants every night at 9 PM. For groups with personalized features (the full group and the personalized group), we also asked to access the user's calendar and location data to be used in recommendations.

A poststudy questionnaire was conducted at the end of the study to evaluate participants' experiences during the first 60 days (Multimedia Appendix 2). First, we provided 3 general close-ended statements to measure participants' overall motivation, satisfaction, and preference with the in-game experience. Participants responded on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). The statements were as follows:

1. I find this kind of application motivating to exercise.
2. I was overall satisfied with this application.
3. I prefer using this type of application for exercise over regular exercises.

We used the Intrinsic Motivation Inventory (IMI; Multimedia Appendix 3) [64] to assess participants' level of intrinsic motivation related to the game experience. Furthermore, we used the European Microsoft Innovation Center (EMIC) recommender system evaluation measurement (Multimedia Appendix 4) [65] to evaluate the quality of our recommended activities. We also included open-ended questions to obtain participants' comments and suggestions to improve the system. By the end of the 60 days, each participant received a Can $10 (US $7.7) gift card as an honorarium to thank them for their participation in the study.

Moreover, we customized the IMI scale to fit the current game context. We did not use *relatedness* and *perceived choice* IMI subscales. *Relatedness* evaluates the experience of doing something with another person, that is, social interactions with a game that can lead to the feelings of relatedness. It is usually used in multiplayer games, which allow for interactions between real players, and was not applicable in our case. *Perceived choice* is often used in situations where a person is given a certain task or activity to complete. In our case, we indicated in the beginning that the users have the full choice of either using or not using our system as well as how to use it. Therefore, this subscale was considered not necessary as the participants were explicitly given full choice. There are different versions of the IMI that have been used in previous studies, which consist of different subscales that are only relevant to their unique context.

Data Analysis

For each question in the poststudy questionnaire including general perception, IMI subscales, and EMIC subscales, a one-way between-group analysis of variance (ANOVA) and *post hoc* Tukey-Kramer Honestly Significant Difference (HSD) test [66] was conducted to analyze the main effects among the 4 groups. ANOVA is commonly used to determine whether there are any statistically significant differences between the means of 3 or more independent groups, whereas the Tukey test provides deeper insights into patterns and comparing specific groups [66]. Parametric tests were selected for conducting the analysis because the samples were drawn independently of each other and the shapes of the distributions were normal. The alpha value was set at .05 for all statistical tests.

For other users' daily log data, such as the number of active users, the number of conversations, the active calories, and the number of useful recommendations, we visualized them along the timeline to see how the pattern differentiated among the 4 groups.

For qualitative data regarding the possible improvement of the system, because our participants' answers were mostly short and concise, we simply categorized them and reported the most commonly mentioned suggestions.

Results

General Information

The participants' self-reported average hours of exercise per week before the study were 4.16 hours with an SD of 2.96 hours, whereas the average hours per week spent playing video games (including PC, console, and mobile games) were 5.44 hours with an SD of 4.13 hours. The self-reported average active hours increased from 4.16 to 4.58 hours after the study.

Participants could interact with the app through their Android phones or watches. Data show that participants read 53.00% (10,270/19,377) of messages on their phones and 47.00% (9107/19,377) of messages on their watches. They tapped 35.81% (1785/4985) of prompted choices on their phones and 63.99% (3190/4985) on their watches. The results suggest that smartwatches were not only effective and more accurate for tracking activity data but also feasible for some simple interactions such as reading messages and tapping a choice from prompts. Participants tended to interact with watches independently when completing simple tasks and switched to

phones when different interactions were necessary (eg, typing messages).

Case Studies

Below, we present 2 case studies as examples to show how our system recommends activities to different participants in a typical week. If there is any more information or any change in the system found during the week, the recommendations adjust accordingly. Both participants were from the full group receiving activity recommendations in the form of a gamified story.

Case Study A

Participant Information

Participant A was a female, 26-year-old student, height 5'8'', weight 61 kg, BMI 20.5 kg/m^2 (normal weight), no serious health issues, and currently taking no medications. Player type: free spirit; fitness color: white. Our system detected that participant A takes the bus to university every Monday, Tuesday, and Thursday and mostly stays at home for the rest of the week. She goes to a group-cycling class once a week, on Friday evenings, for half an hour. According to the GRPAH, adults aged 18 to 64 years were encouraged to perform 300 min of moderate-intensity aerobic physical activity throughout the week for good health benefits [54]. People with exerciser type of white prefer hiking, running, yoga, cardio, and gym strength training. When accessing her calendar, the system found she had 2 dinner reservations on Thursday and Saturday night, both at 6 PM for the coming week.

System-Generated Activity Recommendations

1. Extending the walking distance to bus stops on every school day (both morning and afternoon, overall 45 min of walking per school day).
2. A 30-min walk for non–school days after dinner.
3. A 1-hour home yoga session on Tuesday 7 PM when the user is generally not active.
4. A hiking morning on Saturday in a nearby park.

Player Type–Based Game Features

The player type of free spirit was assigned the game feature of *theme color customization*. Thus, the reward of completing recommended activities for participant A was to unlock different theme colors.

Case Study B

Participant Information

Participant B was a male, 35-year-old, software developer, height 5'11'', weight 75 kg, BMI 23.0 kg/m^2 (normal weight), no serious health issues, and currently taking an over-the-counter pain reliever for his back pain. Player type: achiever; fitness color: red. Our system detected that participant B drives to work every Monday, Tuesday, Thursday, and Friday (15-min drive).

On Wednesday, he works from home. He plays basketball every Wednesday night from 8 PM to 9 PM and every Saturday morning from 9 AM to 11 AM. Exerciser type of reds prefer exercises such as basketball, tennis, racquetball, in-line skating, frisbee, mountain biking, soccer, and skiing. Our system found that participant B was almost as active as recommended by the GRPAH, but the type of activities he performed was limited to basketball.

System-Generated Activity Recommendations

1. A 1-hour tennis or racquetball session on Wednesday night instead of basketball.
2. A daily 15-min walk after work.
3. A 60-min walk (in a nearby park) on Sunday morning.

#### Player Type–Based Game Features

The player type of achiever was assigned the game feature of challenge. Thus, the system provided recommendations to player B in the challenge style.

Overall Motivation and Satisfaction

Figure 5 shows the averages and SDs of the scores for the first 3 general questions assessing participant motivation, satisfaction, and game preference. The asterisk indicates significant results found between groups.

The results show that there were statistically significant differences between groups as determined by one-way ANOVA for overall motivation ($F_{3,36}$=22.49; P<.001), satisfaction ($F_{3,36}$=22.12; P<.001), and preference ($F_{3,36}$=15.0; P<.001). *Post hoc* comparisons using the Tukey HSD test indicated that for all 3 questions, the mean score for the full, personalized, and gamified groups was significantly different from that for the control group, respectively. This means that, in general, both gamification and personalization have positive effects on participants' motivation, satisfaction, and preference, as seen in the groups full, personalized, and gamified compared with the control group. Moreover, for motivation, the mean score for the full group (mean score for full group [MF] 5.8, SD for full group [SDF] 0.79) was significantly different from that of the personalized group (mean score for personalized group [MP] 4.7, SD for personalized group [SDP] 1.5). Statistically significant pairwise comparisons are also marked in Figure 7. This means that gamification can also add more motivation to a personalized fitness recommendation system, as seen between the full group and the personalized group in motivation. It should also be noted that the distribution of the dominant player types across the 4 different groups could have influenced these results, as some player types may have had a stronger preference for gamification or personalization in general. However, the distribution with respect to player types did not seem to be particularly biased (Table 3).

Figure 7. Results for poststudy questions 1, 2, and 3. (a) Overall motivation, (b) overall satisfaction, and (c) overall preference over regular exercise.
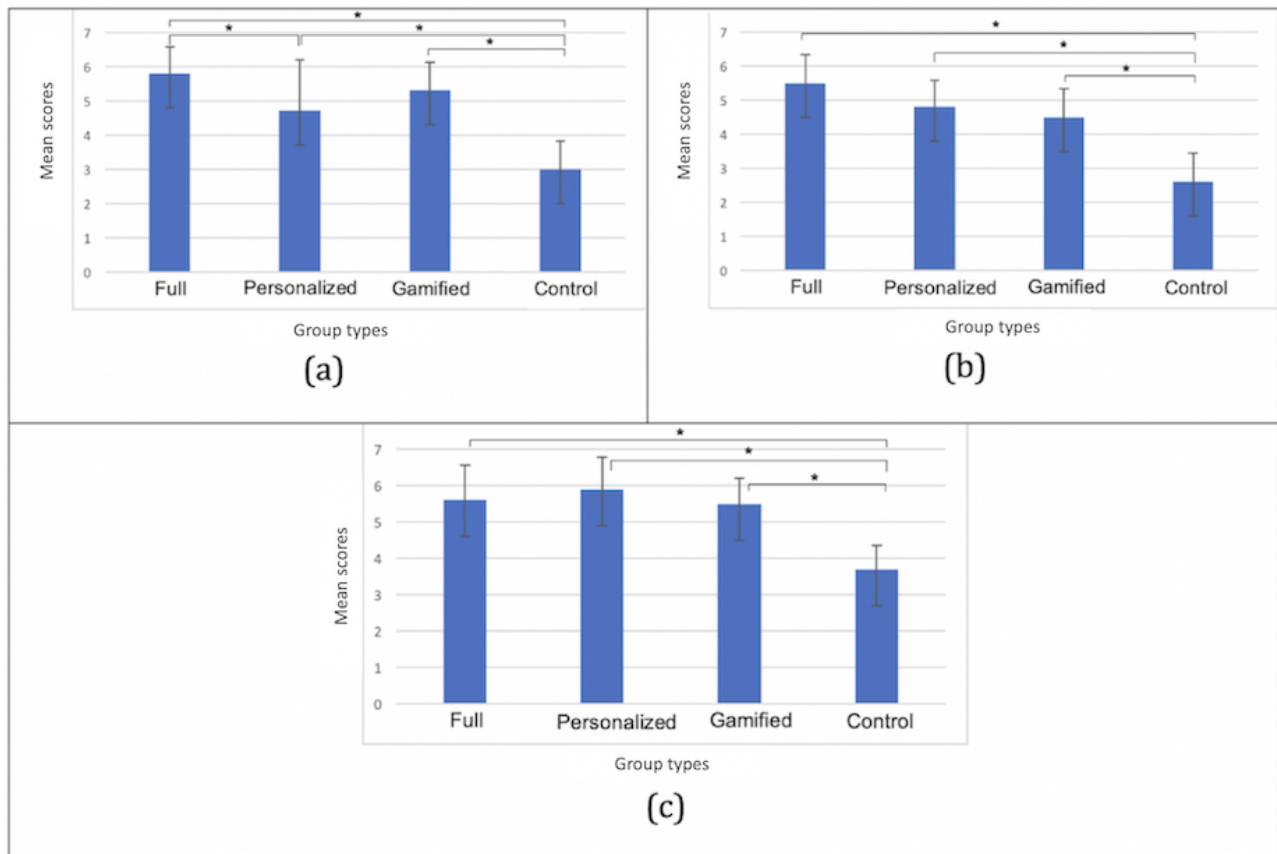

IMI Subscales

Figure 8 shows the average and SDs of the scores for each IMI subscale question. The results show that there were statistically significant differences between groups as determined by a one-way ANOVA for interest or enjoyment ($F_{3,36}$=24.24; P<.001), perceived competence ($F_{3,36}$=4.60; P=.007), effort or importance ($F_{3,36}$=8.01; P<.001), and value or usefulness ($F_{3,36}$=15.90; P<.001).

The Tukey-Kramer HSD test results indicated that for interest or enjoyment, the mean score for the full, personalized, and gamified groups was significantly different from that for the control group. Moreover, the pairwise comparison result showed that MF (MF 5.9, SDF 0.40) was significantly different from the personalized group (MP 5.0, SDP 0.56). For perceived competence, significant differences were found between the full group (MF 5.7, SDF 0.46) and the personalized group (MP 4.8, SDP 0.72) as well as between the full group and the control group (mean score for control group [MC] 5.0, SD for control group [SDC] 0.54). For effort or importance, significant differences were found between the full group (MF 5.8, SDF 0.47) and the gamified group (mean score for gamified group [MG] 4.7, SD for gamified group [SDG] 0.67); between the full group and the control group (MC 4.7, SDC 0.70), the personalized group (MP 5.6, SDP 0.73), and the gamified group; and between the personalized group and the control group. For value or usefulness, significant differences were also found between the full group (MF 5.8, SDF 0.63) and the gamified group (MG 4.8, SDG 0.55); between the full group and the control group (MC 4.6, SDC 0.41), the personalized group (MP 5.7, SDP 0.63), and the gamified group; and between the personalized group and the control group. The pairwise comparison significance is also marked in Figure 8.

Figure 8. Averages and SDs as evaluated by the Intrinsic Motivation Inventory. From top to bottom: (a) interest or enjoyment, (b) perceived competence, (c) effort or importance, (d) pressure or tension, and (e) value or usefulness.
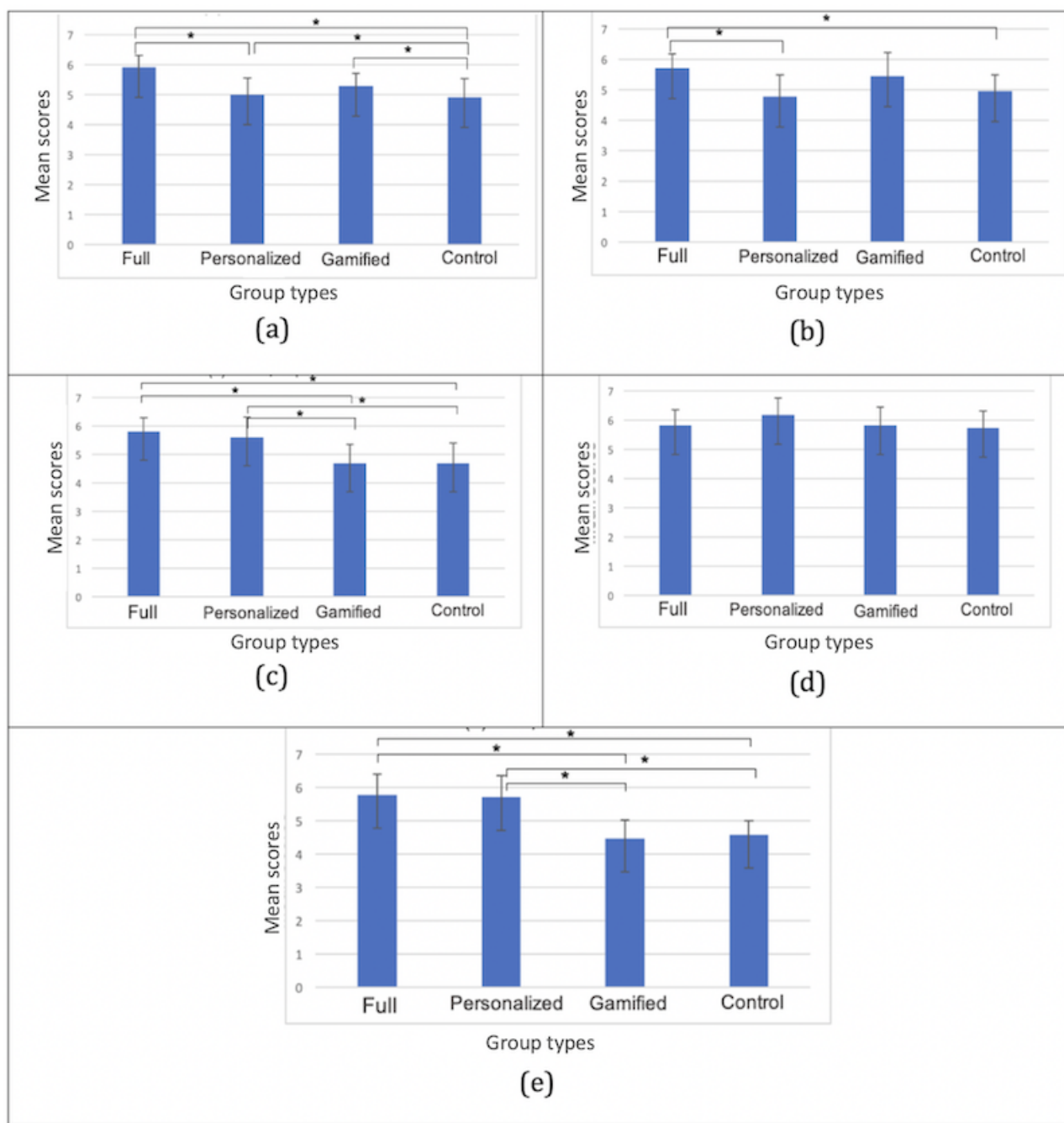
The IMI results indicate that gamifying the exercise increases players' interest in and enjoyment of the personalized recommendation system (significant between the full group and the personalized group in interest or enjoyment). Personalization contributes more toward promoting effort or importance as well as value or usefulness compared with gamification (significant between the personalized group and the gamified group).

EMIC Recommender System Evaluation

Figure 9 shows the averages and SDs of the scores for each EMIC subscale (under perceived recommendation quality, perceived system effectiveness, general trust in technology, and system-specific privacy concerns). The results showed that there were statistically significant differences between groups as determined by a one-way ANOVA for perceived recommendation quality ($F_{3,36}$=108.77; P<.001), perceived system effectiveness ($F_{3,36}$=26.52; P<.001), and system-specific privacy concern ($F_{3,36}$=58.37; P<.001).

Figure 9. Average and SD of European Microsoft Innovation Center recommendation. From top to bottom: (a) perceived recommendation quality, (b) perceived system effectiveness, (c) general trust in technology, and (d) system-specific privacy concern.
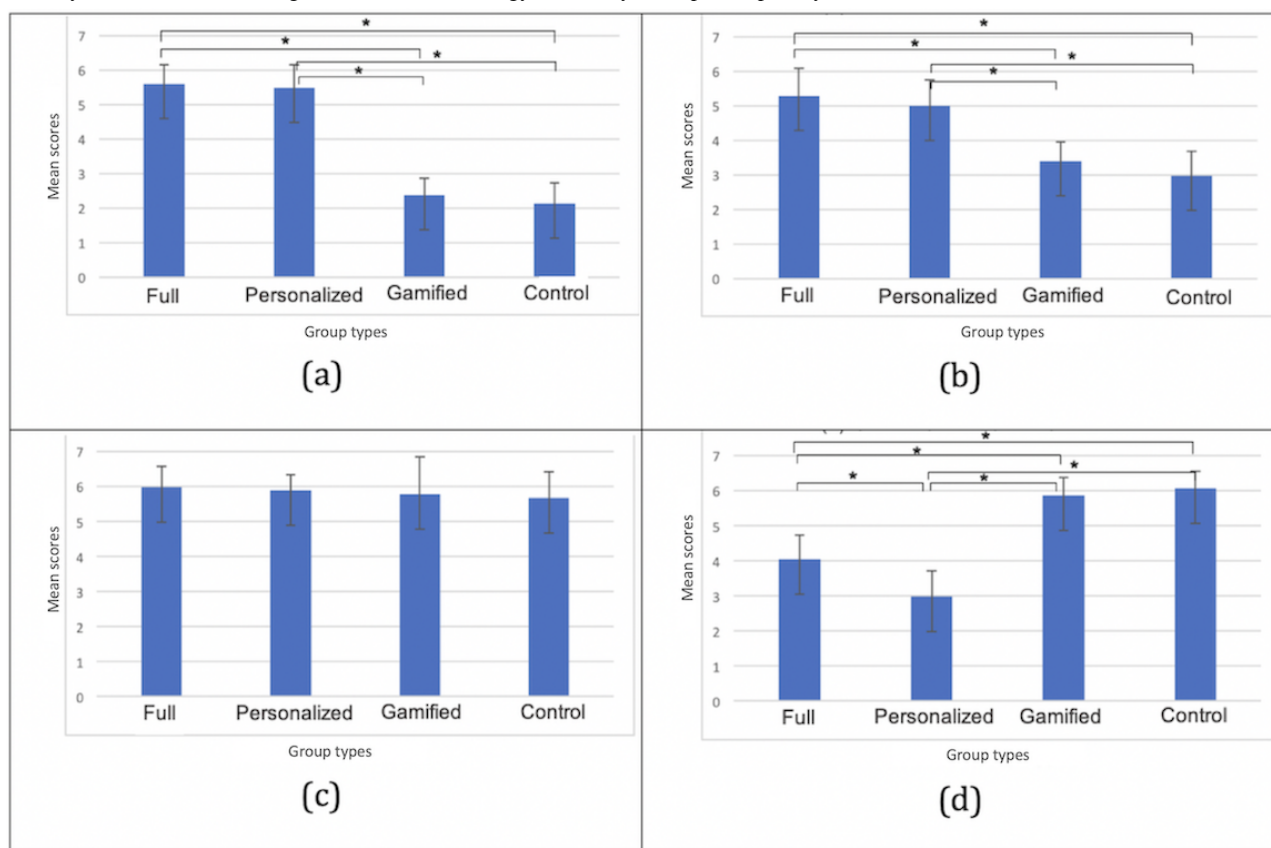
The Tukey-Kramer HSD test results indicated that for both perceived recommendation quality and perceived system effectiveness, the mean scores for the full and personalized groups were significantly different from the gamified and control groups that were not personalized. For system-specific privacy concerns, the mean scores for the full group and the personalized group were also significantly different from the gamified group and the control group because, for nonpersonalized groups, we did not ask to access participants' personal data (except Google Analytics for in-app tracking). Moreover, a significant difference was also found between the full group (MF 4.1, SDF 0.69) and the personalized group (MP 3.0, SDP 0.73). Statistically significant pairwise comparisons are also marked in Figure 9 using asterisks.

Our results suggest that our system is effective in providing daily fitness recommendations to participants (comparing the full group with the gamified group and the personalized group with the control group) with respect to both perceived recommendation quality and perceived system effectiveness. We also found that, as expected, participants were concerned about privacy when the system had a player model and asked for more permissions to access their personal data (comparing the full or personalized and gamified or control groups). On the other hand, gamification reduced some of the concerns (significant difference found between the full and the personalized groups). Note that for the system-specific privacy question, a higher score indicates less concern. Privacy concerns are important yet beyond the scope of this work. Yet, we believe that the noticed effect of gamification can be of value in future research and design.

Daily Statistical Data

As mentioned earlier, we used Google Analytics API to track participants' comprehensive in-app behavior data and we used Google Fit API to track user daily activity data, including steps and calories burned. Figure 10 shows some daily statistical data: the number of active participants of all 4 groups during the 60 days of study (a), the daily total number of conversations sent to the system (b), the daily average active calories burned excluding basal metabolism (c), and the daily number of self-reported useful recommendations (d).

Figure 10. Daily statistical data showing patterns of daily active users, conversations, or active calories and percentage of useful recommendations along the timeline.

(a) Daily active users

(b) Daily conversations

(c) Daily active calories

(d) Daily percentage of useful recommendations
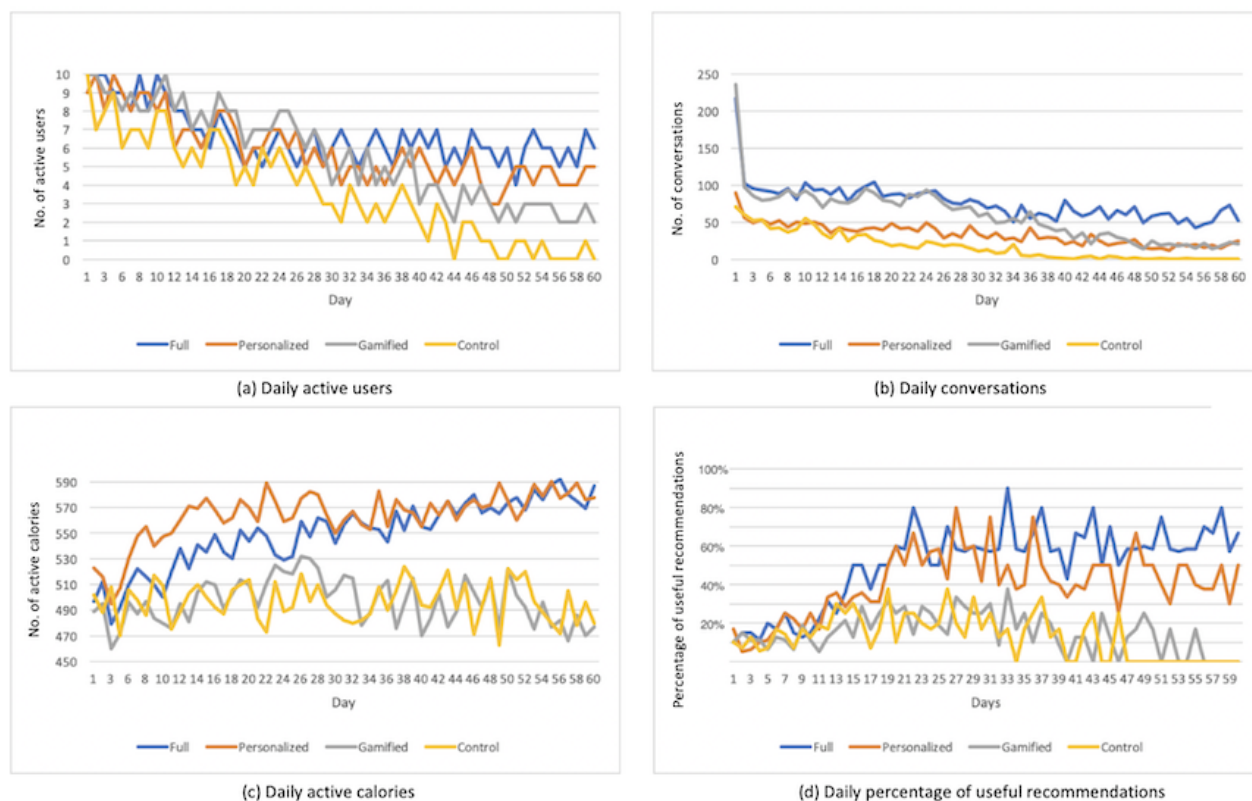
From Figure 8, we can see that for daily active users (a) and daily conversations (b), there is an overall descent in trends appearing as time grows for all 4 groups. Among them, the full group maintained a relatively higher value compared with the other 3 groups and participants in the full, personalized, and gamified groups interacted with the system more than the control group (Figure 10). With respect to the daily active use and daily conversations (Figure 10), when comparing the personalized group and the gamified group, we can see that the value of the gamified group was higher than that of the personalized group in the early phase of the study but was surpassed by the personalized group in the late phase of the experiment (around 35-40 days). These results indicate that both personalization and gamification could have a positive impact on promoting participants' engagement with the system. However, although gamification could bring more interactions in the short term (within 1 month), personalization could lead to a more sustained engagement (over a longer time). Note that Figure 10 shows that the control group was not active during the last week of the study. This only indicates that they did not open the app but they still received recommendations as usual (pop-up notifications). Physical activity data were also collected from the Google Fit API without opening the app.

For active calories (Figure 10), we can see slight ascent trends for both the full group and the personalized group and flat trends for the gamified group and the control group. The full group began with a lower average calorie burden compared with the personalized group and then showed an almost equal value near the end. These results indicate that personalization could have a positive impact on promoting actual physical activity, whereas exclusive gamification may not. Adding gamified elements to personalized recommendations in the earlier phase (when the player model was not well established yet and the recommendation quality was not steady enough) may negatively affect the amount of physical activity people performed, which requires further research. Note that the active calorie measures the calories burned during fitness activities. Basal metabolic parameters were excluded.

For the percentage of useful recommendations, Figure 10 (calculated by the daily number of *useful* replies divided by daily active users), the percentage of the full group and the personalized group increased in the first half of the study and then remained flat, with the full group remaining slightly higher than the personalized group. The increase in the full group and the personalized group can be attributed to the continuously updating player model that will improve recommendations over time. The gamified group and the control group (without player model) showed descending trends approaching zero. The results suggest that our system is able to generate useful fitness recommendations by using a player model, and participants considered the recommendation more useful when gamification elements were added.

Player Types

In this study, we did not find any significant difference in terms of different player or exerciser types. Although we had a limited sample size for conducting a meaningful statistical analysis, there were still some interesting findings worth mentioning, which may help inspire future research in this area. Table 4 shows the distribution of the combinations of player and exerciser types of our participants.

From Table 4, we can see that certain player types and exerciser types were highly related. For example, we have 5 participants in total with the exerciser type of silver, with 4 of them belonging to the player type of free spirit. Similar relations are shown between the player type of socializer and the exerciser type of purples. This indicates that users' preferences toward game elements and exercise types may be linked. This idea could be used to further improve the personalization of the exercise and game experience but requires further research with a larger sample size.

Table 4. Distribution of the combination of player or exerciser type (N=40).

The 8 colors	Achiever	Player	Socializer	Philanthropist	Disruptor	Free spirit
Blue	3	3	0	2	0	0
Gold	1	1	1	2	0	0
White	2	1	0	1	0	2
Purple	0	0	2	0	0	0
Green	0	1	3	2	0	1
Red	2	0	0	1	1	0
Saffron	1	0	0	0	1	1
Silver	0	1	0	0	0	4

Figure 11 shows the overall motivation for participants belonging to different player types. Although we were not able to run a valid statistical analysis based on the small sample sizes, we saw that the player type of socializer and disruptor experienced lower overall motivation compared with the other 4 player types. This may indicate that the game features and experience we provided to the player type of socializers and disruptors had more room for improvement.

Figure 11. Overall motivation for different player types.
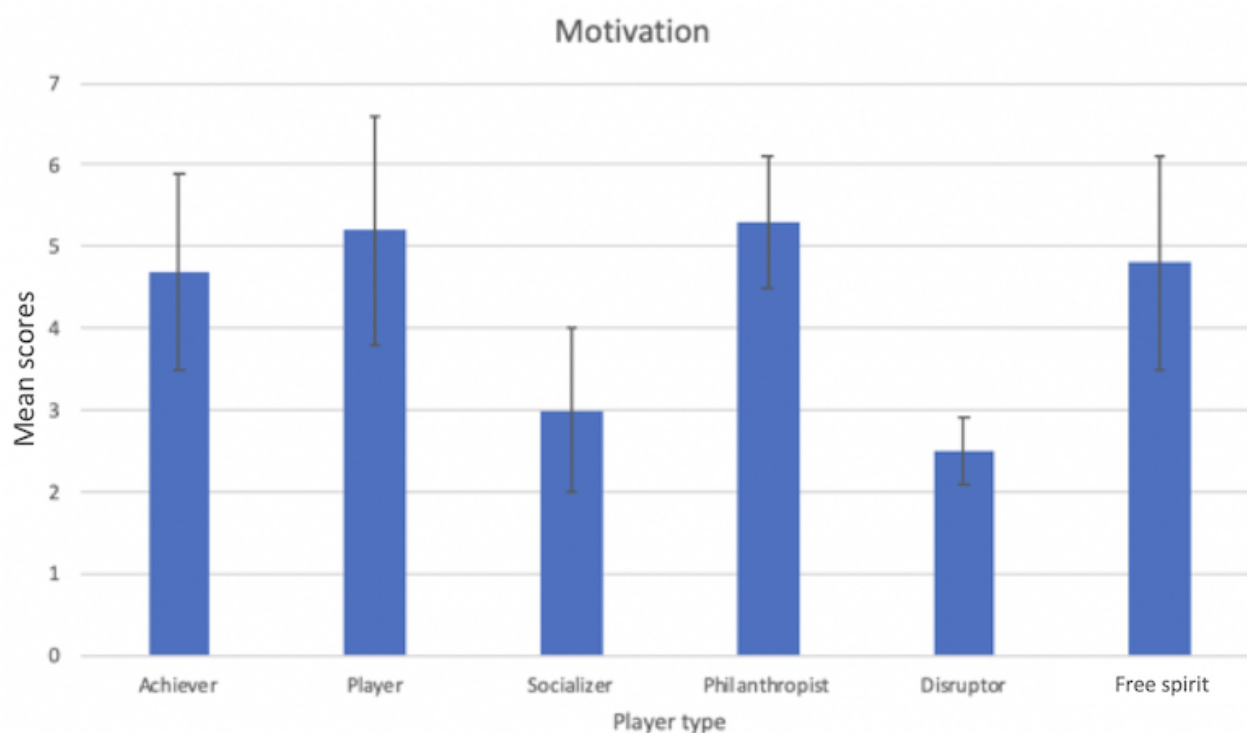
Figure 12 shows the average active calories burned for participants belonging to different exerciser types. For the same reason of small sample sizes, we could not run a valid statistical analysis. However, we found that the fitness colors of whites and greens were relatively more active during the study. We checked their motivation as well as their self-reported recommendation quality and found that both were at the same level as the other 6 exerciser types. When looking at the 8 Colors of Fitness activity suggestions (Multimedia Appendix 5), we found that the activity of hiking was the main variable that may lead to the result and it was only recommended for the exerciser type of greens and whites. This indicates that hiking might be an effective activity that makes people consume more calories, which could be further investigated.

Figure 12. Average active calories for different exerciser types.
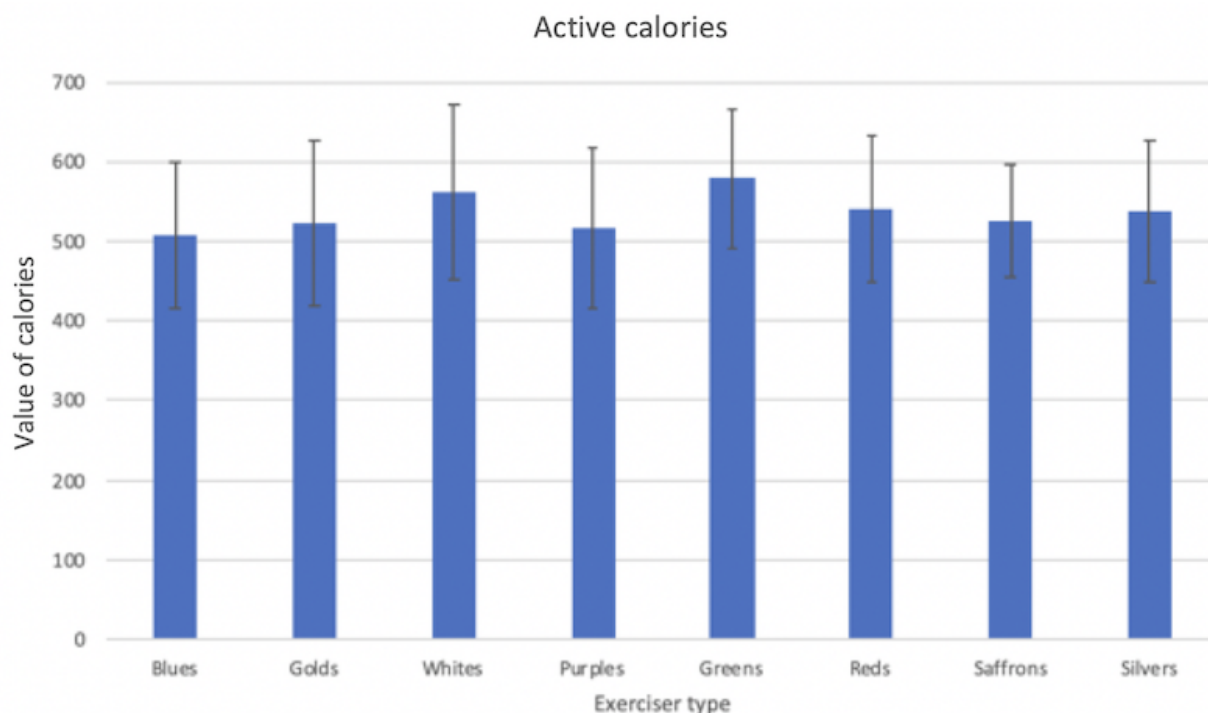
Qualitative Results

One open-ended question was asked of each participant at the end of the poststudy questionnaire to collect their general feedback (Multimedia Appendix 2). We received many comments and suggestions on how to improve our system, which mainly focused on 5 aspects as shown in Figure 13. It shows that a customized storyline was the most requested feature, followed by multiplayer mode, more quality recommendations, a feature for setting and tracking fitness goals, and more location-based features. These feedback laid the foundation for planning our future work in this project.

Figure 13. Number of main suggestions received from open-ended questions for improving our system.
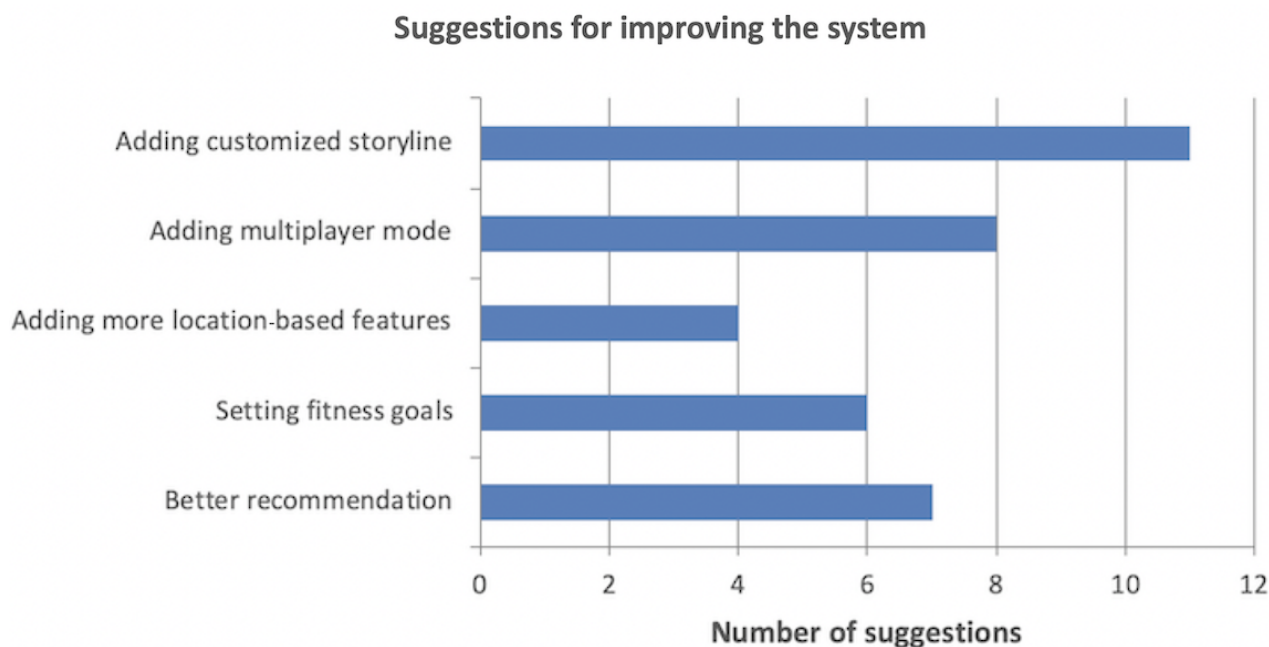
Discussion

Overall, in this 60-day user study, we verified our hypotheses that (1) it is feasible to generate personalized exercise recommendations with player modeling and (2) the combination of player modeling and gamification could enhance users' engagement with the system as well as promote actual physical activity. Specifically, gamification was found to promote engagement, but only in the short term, as seen in the gamified group where the members were engaged early on. However, as

the experiment moved on, the trend changed and the personalized group became more engaged. This can be attributed to the player modeling aspect in that it requires time to get to a minimum level of precision in reflecting a player's characteristics before it can offer reasonable recommendations. Player modeling helped sustain the activity level in the long term. This suggests that activity recommendation based on player modeling can be an effective and promising approach for creating personalized fitness experiences over longer periods, whereas gamification can help attract the users and create the initial interest.

Our research was motivated by the need to keep players engaged and motivated in exergames. We were inspired by previous work that suggested a more player-centric and personalized approach to game design and gamification [6,17,28,67] to increase player engagement and the overall effectiveness of the intervention. We extended these ideas to exergames, combining them with the notion of real-time activity tracking and recommendation as suggested by others [22,47-49] to develop a new theoretical dynamic and individual-level model that brings together various game elements that can help solve the player retention problem. The presented results have direct implications for the design of fitness assistants and potentially other recommender systems.

Gamification Is Good but Not Enough!

Previous work by authors and other researchers has shown the potential value of gamification to increase engagement, but they have also highlighted the issue of retention. Players tend to leave the game once it is well experienced. Although adding new features can be a reasonable way of keeping participants engaged, it is difficult and costly to implement because of constant designing and upgradation. The ability to understand participants and their dynamic life and provide gameplay features that match the participants' activities can be a way to introduce change and novelty when maintaining the development cost under control.

Player Modeling: Personalization Versus Categorization

The idea of categorizing participants to provide them with customized service is appealing but ignores individual differences, which are often significant. The availability of personal data, as a result of various methods of collecting information, suggests that the participants can be understood as individuals and not members of a category. This true personalization allows a new level of customization that will potentially offer participants a much more appealing and effective experience. Our results show the potential relevance of this idea to the field of fitness assistants. The more we understand the user, the more personalized our recommendations will be, which will, in turn, result in more effective recommendations. Developing a comprehensive model that involves various user characteristics (from personality type to daily routines) can help understand the user properly.

Furthermore, the idea of personalization versus categorization is also related to differences in player types and player traits. Although earlier works have attempted to classify players into single types (eg, Bartle [68] and the BrainHex model [36]), more recently, researchers have examined the effectiveness of trait-oriented models for understanding player choices in games [69-71]. Trait-oriented models are preferred in recent studies because an individual is rarely motivated by a single factor and because of their applicability to game user research in that they aim to characterize players using a set of scores rather than categorizing players into a single type. In this study, we decided to use the dominant player type as evaluated by the Hexad model rather than considering all 6 scores because we wanted to control the variable by adding only 1 additional element to each user; therefore, we could make sure it is the gamification itself that affected the engagement, without interfering with the amount of it. A future study can explore the effects of considering the full range of scores.

Adaptive and Continuous Modeling

Although many games and other applications rely on a certain user model, in most cases, this is done as a one-time static decision assigning the user to a certain group. Our study shows the value of not only having a more comprehensive personal model but also allowing it to evolve and adapt using ongoing data from the user. This constantly tunes the model and makes recommendations more effective. Using such adaptive and dynamic models can enhance the performance of such applications, and we recommend that designers consider it when possible.

24/7 Recommendation

Fitness and health are not limited to the gyms. Being active is a lifestyle; therefore, activity recommendations should not be limited to a particular time. In the absence of a dedicated personal trainer, an intelligent fitness assistant equipped with a detailed player model can offer 24/7 recommendations for being active that considers various user contexts. Our results show the potential value of this approach, which can be improved with more comprehensive personal data and a better database of activities and gameplay features. Although our system provided all-day and continuous modeling and recommendation, it is worth noting that the participants did not wear the activity trackers during sleep and we did not track any sleeping activities. As such, although the system was able to perform nonstop, in practice, it was paused during sleep times (night or day).

Limitations

There were certain limitations in the proposed system and the performed study, some mentioned by the participants, which we believe were not critical enough to significantly affect the findings but are still worth noting and improving in future work.

We relied on a simple game that we designed ourselves with a simple story or gameplay. This may have negatively affected the players' attraction and engagement. The game could be designed through a more rigorous process or we could somehow allow customization and choice or potentially use another existing game. There was also no multiplayer option, which ignores the social aspects of gaming and active lifestyle and could negatively affect the level of user engagement. When designing different gamified features for different types of

players, we assigned only 1 game element to each type of player. This may not be adequate for targeting individual participants.

The 8 Colors of Fitness system (Multimedia Appendix 5) [44] was used as a model to suggest activities. This system was used because the research group did not find any other alternatives and needed to rely on a fairly acceptable method. This system is by no means ideal and has its own limitations. It can be replaced with any other method, such as other models, an interactive trainer, or a trained expert system.

We used the Android Activity Recognition API for activity tracking and prediction in this work. This API is only able to recognize 6 simple physical activities. For more complex daily activities, we required manual labeling from participants within the conversation. This may bring complexity to the participants. We also only used Android Wear participants and limited each group to 10 members, which may not be adequate. We were also aware that the age range of our participants was relatively narrow. Most of our participants in this study were young adults; hence, our results may not apply to older adults. Furthermore, comparing active calories burned as an absolute value could have negatively influenced the reliability of the results because of potential confounding variables such as gender, weight, and height.

The language of our questions could be improved by being more neutral and consistent. For example, we occasionally used *task* to refer to the app instead of the more common term *game*, or for EMIC, we used *items* and *activities* for the same purpose. Although these terms could have caused some confusion, which we will improve in the future, we did not receive any negative feedback and do not believe that the inconsistencies significantly affected our findings.

Conclusions

In this paper, we proposed a system for personalized fitness assistants using gamification and continuous player modeling and reported on a long-term study that investigates the effectiveness of our proposed system. Our findings show that it is possible to provide personalized activity recommendations by continuously updating a player model based on activity tracking. Our study also shows the positive effect of this modeling and gamification on user engagement and overall activity. These findings can be used to inform the design of personalized and gamified recommender systems in health and fitness and potentially other apps, as they highlight the role of an adaptive model and gamification as long-term and short-term factors, respectively. This research opens opportunities for future work, especially in the area of exploring more gameplay features, adding a personalized storyline, multiplayer gamification, better activity recognition, suggestion models, and evaluation with a larger and more diverse sample.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Prestudy questionnaire for collecting general information.
[DOCX File , 813 KB - games_v8i4e19968_app1.docx]

Multimedia Appendix 2
Poststudy questionnaire for collecting general feedback. EMIC: European Microsoft Innovation Center; IMI: Intrinsic Motivation Inventory.
[DOCX File , 13 KB - games_v8i4e19968_app2.docx]

Multimedia Appendix 3
Intrinsic Motivation Inventory for evaluating the level of enjoyment related to the game experience.
[DOCX File , 14 KB - games_v8i4e19968_app3.docx]

Multimedia Appendix 4
European Microsoft Innovation Center recommender system evaluation measurement tool.
[DOCX File , 14 KB - games_v8i4e19968_app4.docx]

Multimedia Appendix 5
Eight Colors of Fitness activity suggestions.
[DOCX File , 13 KB - games_v8i4e19968_app5.docx]

References

1. Wilmot EG, Edwardson CL, Achana FA, Davies MJ, Gorely T, Gray LJ, et al. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. Diabetologia 2012 Nov;55(11):2895-2905. [doi: 10.1007/s00125-012-2677-z] [Medline: 22890825]

XSL•FO

RenderX

2. Hagen K, Chorianopoulos K, Wang AI, Jaccheri L, Weie A. Gameplay as Exercise. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Presented at: CHI EA'16; October 21-24, 2016; San Jose, CA. [doi: 10.1145/2851581.2892515]

3. Dharia S, Eirinaki M, Jain V, Patel J, Varlamis I, Vora J, et al. Social recommendations for personalized fitness assistance. Pers Ubiquit Comput 2017 Jun 29;22(2):245-257. [doi: 10.1007/s00779-017-1039-8]

4. Ananthanarayan A, Siek KA. Persuasive Wearable Technology Design for Health and Wellness. In: 6th International Conference on Pervasive Computing Technologies for Healthcare and Workshops. 2012 Presented at: PervasiveHealth'12; May 2-7, 2012; San Diego, CA. [doi: 10.4108/icst.pervasivehealth.2012.248694]

5. McGonigal J. Reality is broken: why games make us better and how they can change the world. Choice Rev Online 2012 Jul 1;49(11):49-6095. [doi: 10.5860/choice.49-6095]

6. Nacke LE, Deterding S. The maturing of gamification research. Comput Hum Behav 2017 Jun;71:450-454. [doi: 10.1016/j.chb.2016.11.062]

7. Prensky M. Digital game-based learning. Comput Entertain 2003 Oct;1(1):21. [doi: 10.1145/950566.950596]

8. Fogg BJ. Persuasive technologies. Commun ACM 1999;2(4):26-29. [doi: 10.1007/978-3-319-07127-5_11]

9. Papastergiou M. Exploring the potential of computer and video games for health and physical education: a literature review. Comput Educ 2009 Nov;53(3):603-622. [doi: 10.1016/j.compedu.2009.04.001]

10. Hamari J. Do badges increase user activity? A field experiment on the effects of gamification. Comput Hum Behav 2017 Jun;71:469-478. [doi: 10.1016/j.chb.2015.03.036]

11. Hamari J, Koivisto J, Sarsa H. Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. In: 47th Hawaii International Conference on System Sciences. 2014 Presented at: CSS'14; January 6-9, 2014; Waikoloa. [doi: 10.1109/HICSS.2014.377]

12. Baranowski T, Buday R, Thompson DI, Baranowski J. Playing for real: video games and stories for health-related behavior change. Am J Prev Med 2008 Jan;34(1):74-82. [doi: 10.1016/j.amepre.2007.09.027] [Medline: 18083454]

13. Consolvo S, Klasnja P, McDonald DW, Landay JA. Goal-setting Considerations for Persuasive Technologies That Encourage Physical Activity. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Presented at: Persuasive'09; April 11-16, 2009; California, Claremont, USA. [doi: 10.1145/1541948.1541960]

14. Orji R, Mandryk RL, Vassileva J, Gerling KM. Tailoring Persuasive Health Games to Gamer Type. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI '13; March 3-8, 2013; Paris, France. [doi: 10.1145/2470654.2481341]

15. Marquet O, Alberico C, Adlakha D, Hipp JA. Examining motivations to play pokémon go and their influence on perceived outcomes and physical activity. JMIR Serious Games 2017 Oct 24;5(4):e21. [doi: 10.2196/games.8048] [Medline: 29066423]

16. Orji R, Mandryk RL, Vassileva J. Improving the efficacy of games for change using personalization models. ACM Trans Comput-Hum Interact 2017 Nov 13;24(5):1-22. [doi: 10.1145/3119929]

17. Orji R, Tondello GF, Nacke LE. Personalizing Persuasive Strategies in Gameful Systems to Gamification User Types. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI'18; April 3-5, 2018; Montreal, Canada. [doi: 10.1145/3173574]

18. Charles D, McNeill M, McAlister M, Black M, Adrian M, Stringer K, et al. Player-Centred Game Design: Player Modelling and Adaptive Digital Games. In: Proceedings of DiGRA 2005 Conference: Changing Views – Worlds in Play. Digital Games Research Association. 2005 Presented at: DiGRA'05; March 4-8, 2005; Vancouver, Canada URL: https://www.researchgate.net/publication/228636408_Player-centred_game_design_Player_modelling_and_adaptive_digital_games

19. Yannakakis GN, Spronck P, Loiacono D, Andre E. Player Modeling. Dagstuhl Publishing. 2013. URL: http://yannakakis.net/wp-content/uploads/2013/08/pm_submitted_final.pdf [accessed 2020-11-10]

20. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. JMIR Serious Games 2014 Aug 4;2(2):e9 [FREE Full text] [doi: 10.2196/games.3413] [Medline: 25654660]

21. Campbell T, Ngo B, Fogarty J. Game Design Principles in Everyday Fitness Applications. In: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work. 2008 Presented at: CSCW'08; March 1-6, 2008; San Diego, CA. [doi: 10.1145/1460563.1460603]

22. Rabbi M, Aung MH, Zhang M, Choudhury T. MyBehavior: Automatic Personalized Health Feedback From User Behaviors and Preferences Using Smartphones. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: UbiComp'15; April 1-4, 2015; Osaka, USA. [doi: 10.1145/2750858.2805840]

23. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]

24. Mekler ED, Brühlmann F, Tuch AN, Opwis K. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. Comput Hum Behav 2017 Jun;71:525-534. [doi: 10.1016/j.chb.2015.08.048]

25. Loria E. A Framework to Infer Player Experience and Inform Customized Content Generation in Gameful Systems. In: Extended Abstracts: Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. 2019 Presented at: CHI PLAY'19; August 21-29, 2019; Barcelona, Spain. [doi: 10.1145/3341215.3356330]

26.   Böckle M, Micheel I, Bick M, Novak J. A Design Framework for Adaptive Gamification Applications. In: Hawaii
       International Conference on System Sciences. 2018 Presented at: ICSS'18; January 6-9, 2018; Hawaii, USA. [doi:
       10.24251/hicss.2018.151]

27.   Koivisto J, Hamari J. The rise of motivational information systems: a review of gamification research. Int J Inf Manag 2019
       Apr;45:191-210. [doi: 10.1016/j.ijinfomgt.2018.10.013]

28.   Bertran FA, Segura EM, Duval J, Isbister K. Chasing Play Potentials: Towards an Increasingly Situated and Emergent
       Approach to Everyday Play Design. In: Proceedings of the 2019 on Designing Interactive Systems Conference. 2019
       Presented at: DIS'19; February 7-10, 2019; San Diego, CA. [doi: 10.1145/3322276.3322325]

29.   Macvean A, Robertson J. Understanding Exergame Users' Physical Activity, Motivation and Behavior Over Time. In:
       Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI'13; March 3-9,
       2013; Paris, France. [doi: 10.1145/2470654.2466163]

30.   Busch M, Mattheiss E, Hochleitner W, Hochleitner C, Lankes M, Fröhlich P, et al. Using player type models for personalized
       game design – an empirical investigation. Interact Design 2016;12(44):145-163. [doi: 10.1145/2851581.2892399]

31.   Velazquez A, Martínez-García AI, Favela J, Ochoa SF. Adaptive exergames to support active aging: an action research
       study. Pervasive Mob Comput 2017 Jan;34:60-78. [doi: 10.1016/j.pmcj.2016.09.002]

32.   Yoo S, Heywood T, Tang LM, Kummerfeld B, Kay J. Towards a Long Term Model of Virtual Reality Exergame Exertion.
       In: Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization. 2017 Presented at: UMAP'17;
       April 6-9, 2017; Bratislava, UK. [doi: 10.1145/3079628.3079679]

33.   Farič N, Yorke E, Varnes L, Newby K, Potts HW, Smith L, et al. Younger adolescents' perceptions of physical activity,
       exergaming, and virtual reality: qualitative intervention development study. JMIR Serious Games 2019 Jun 17;7(2):e11960
       [FREE Full text] [doi: 10.2196/11960] [Medline: 31210135]

34.   Malone LA, Padalabalanarayanan S, McCroskey J, Thirumalai M. Assessment of active video gaming using adapted
       controllers by individuals with physical disabilities: a protocol. JMIR Res Protoc 2017 Jun 16;6(6):e116 [FREE Full text]
       [doi: 10.2196/resprot.7621] [Medline: 28623186]

35.   Wiemeyer J, Deutsch J, Malone LA, Rowland JL, Swartz MC, Xiong J, et al. Recommendations for the optimal design of
       exergame interventions for persons with disabilities: challenges, best practices, and future research. Games Health J 2015
       Feb;4(1):58-62 [FREE Full text] [doi: 10.1089/g4h.2014.0078] [Medline: 26181682]

36.   Richard B. Hearts, clubs, diamonds, spades: players who suit MUDs. J MUD Res 1996 [FREE Full text]

37.   Deterding S, Dixon D, Khaled R, Nacke L. From Game Design Elements to Gamefulness: Defining. In: Proceedings of
       the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at:
       MindTrek'11; April 9-11, 2011; Tampere, UK. [doi: 10.1145/2181037.2181040]

38.   Nacke LE, Bateman C, Mandryk RL. BrainHex: a neurobiological gamer typology survey. Entertain Comput 2014
       Jan;5(1):55-62. [doi: 10.1016/j.entcom.2013.06.002]

39.   Tondello GF, Wehbe RR, Diamond L, Busch M, Marczewski A, Nacke LE. The Gamification User Types Hexad Scale.
       In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play. 2016 Presented at: CHI PLAY'16;
       March 2-8, 2016; Austin, UK. [doi: 10.1145/2967934.2968082]

40.   Wiemeyer J, Nacke L, Moser C, Mueller FF. Player experience. In: Serious Games. New York, USA: Springer; 2016.

41.   Cote J, Dreiss A, Clobert J. Social personality trait and fitness. Proc R Soc B 2008 Aug 26;275(1653):2851-2858. [doi:
       10.1098/rspb.2008.0783]

42.   Courneya KS, Hellsten LM. Personality correlates of exercise behavior, motives, barriers and preferences: an application
       of the five-factor model. Pers Individ Dif 1998 May;24(5):625-633. [doi: 10.1016/s0191-8869(97)00231-6]

43.   Rhodes RE, Smith NE. Personality correlates of physical activity: a review and meta-analysis. Br J Sports Med 2006
       Dec;40(12):958-965 [FREE Full text] [doi: 10.1136/bjsm.2006.028860] [Medline: 17124108]

44.   Brue S. The 8 Colors of Fitness. Corte Madera, CA: Oakledge Press; 2008.

45.   Myers IB, McCaulley MH, Most R. MBTI Manual: A Guide to the Development and Use of the Myers - Briggs Type
       Indicator. London, UK: Pyschologists Press; 1998.

46.   Lara OD, Labrador MA. A survey on human activity recognition using wearable sensors. IEEE Commun Surv Tutorials
       2013;15(3):1192-1209. [doi: 10.1109/surv.2012.110112.00192]

47.   Guo X, Liu J, Chen Y. FitCoach: virtual fitness coach empowered by wearable mobile devices. In: IEEE Conference on
       Computer Communications. 2017 Presented at: INFOCOM'17; March 3-10, 2017; Atlanta, USA. [doi:
       10.1109/infocom.2017.8057208]

48.   He Q, Agu E, Strong D, Tulu B. RecFit: a Context-aware System for Recommending Physical Activities. In: Proceedings
       of the 1st Workshop on Mobile Medical Applications. 2014 Presented at: MMA'14; November 1-6, 2014; Memphis, USA.
       [doi: 10.1145/2676431.2676439]

49.   Broekhuizen K, van Poppel MN, Koppes LL, Brug J, van Mechelen W. A tailored lifestyle intervention to reduce the
       cardiovascular disease risk of individuals with Familial Hypercholesterolemia (FH): design of the PRO-FIT randomised
       controlled trial. BMC Public Health 2010 Feb 15;10(1):69 [FREE Full text] [doi: 10.1186/1471-2458-10-69] [Medline:
       20156339]

XSL•FO

RenderX

50. Mittal R, Sinha V. A Personalized Time-bound Activity Recommendation System. In: 7th Annual Computing and Communication Workshop and Conference. 2017 Presented at: CCWC'17; April 4-9, 2017; Las Vegas, USA. [doi: 10.1109/ccwc.2017.7868485]

51. Ni J, Muhlstein L, McAuley J. Modeling Heart Rate and Activity Data for Personalized Fitness Recommendation. In: The World Wide Web Conference. 2019 Presented at: WWW'19; May 3-8, 2019; San Francisco, CA. [doi: 10.1145/3308558.3313643]

52. Ghanvatkar S, Kankanhalli A, Rajan V. User models for personalized physical activity interventions: scoping review. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11098. [doi: 10.2196/11098] [Medline: 30664474]

53. Zhao Z, Arya A, Whitehead A, Chan G, Etemad SA. Keeping Users Engaged through Feature Updates: A Long-Term Study of Using Wearable-Based Exergames. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI'17; April 3-9, 2017; Denver, CO. [doi: 10.1145/3025453.3025982]

54. Global Recommendations on Physical Activity for Health. Geneva, Switzerland: World Health Organization; 2010.

55. Activity Recognition Client. URL: https://developers.google.com/android/reference/com/google/android/gms/location/ActivityRecognitionClient [accessed 2020-02-11]

56. Google Fit. URL: https://www.google.com/fit/ [accessed 2019-09-11]

57. Wit.ai. URL: https://wit.ai/ [accessed 2019-07-21]

58. Weather API. URL: http://weather.com [accessed 2019-08-01]

59. The Android Wear Forum. URL: https://forums.androidcentral.com/android-wear-wear-os/ [accessed 2018-10-23]

60. Gamified UK User Type Test. URL: https://gamified.uk/UserTypeTest2016/user-type-test.php#.X6XHa2T0lUM [accessed 2019-02-21]

61. The 8 Colors Quiz. URL: http://the8colors.com/quizw.php [accessed 2019-12-02]

62. Visual Studio App Center. Microsoft. URL: https://appcenter.ms/

63. Google Analytics. URL: https://analytics.google.com/analytics/web/ [accessed 2018-01-13]

64. Deci EL, Ryan RM. Intrinsic Motivation Inventory (IMI). Scale Description. 2003. URL: https://selfdeterminationtheory.org/intrinsic-motivation-inventory/ [accessed 2020-11-10]

65. Knijnenburg BP, Willemsen MC, Gantner Z, Soncu H, Newell C. Explaining the user experience of recommender systems. User Model User-Adap Inter 2012 Mar 10;22(4-5):441-504. [doi: 10.1007/s11257-011-9118-4]

66. Jaccard J, Becker MA, Wood G. Pairwise multiple comparison procedures: a review. Psychol Bull 1984:- epub ahead of print. [doi: 10.4135/9781412985178.n2]

67. Nicholson S. A RECIPE for meaningful gamification. In: Gamification in Education and Business. New York, USA: Springer; 2014.

68. Mueller FF, Edge D, Vetere F, Gibbs MR, Agamanolis S, Bongers B, et al. Designing Sports: a Framework for Exertion Games. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI'11; April 7-9, 2011; Vancouver, Canada. [doi: 10.1145/1978942.1979330]

69. Fortes Tondello G, Valtchanov D, Reetz A, Wehbe RR, Orji R, Nacke LE. Towards a trait model of video game preferences. Int J Hum-Comput Interact 2018 Apr 19;34(8):732-748. [doi: 10.1080/10447318.2018.1461765]

70. Tondello GF, Arrambide K, Ribeiro G, Cen AJ, Nacke LE. 'I don't fit into a single type': a trait model and scale of game playing preferences. In: Human-Computer Interaction. 2019 Presented at: INTERACT'19; June 2-7, 2019; Paphos, UK. [doi: 10.1007/978-3-030-29384-0_23]

71. Tondello GF, Nacke LE. Player Characteristics and Video Game Preferences. In: Proceedings of the Annual Symposium on Computer-Human Interaction in Play. 2019 Presented at: CHI PLAY'19; July 6-11, 2019; Barcelona, Spain. [doi: 10.1145/3311350.3347185]

Abbreviations

ANOVA: analysis of variance
API: application programming interface
EMIC: European Microsoft Innovation Center
GRPAH: Global Recommendation on Physical Activity for Health
HSD: honestly significant difference
IMI: Intrinsic Motivation Inventory
MBTI: Myers-Briggs–Type Indicator
MC: mean score for control group
MF: mean score for full group
MG: mean score for gamified group
MP: mean score for personalized group
SDC: SD for control group
SDF: SD for full group
SDG: SD for gamified group

XSL•FO
RenderX

SDP: SD for personalized group
UI: user interface
VR: virtual reality

XSL•FO
RenderX

Original Paper

Cognitive Training Using Fully Immersive, Enriched Environment Virtual Reality for Patients With Mild Cognitive Impairment and Mild Dementia: Feasibility and Usability Study

Seo Jung Yun[1], MS; Min-Gu Kang[1], MD; Dongseok Yang[2], PhD; Younggeun Choi[2], PhD; Heejae Kim[1], MD; Byung-Mo Oh[1], MD, PhD; Han Gil Seo[1], MD, PhD

[1]Department of Rehabilitation Medicine, Seoul National University Hospital, Seoul, Republic of Korea
[2]Department of Computer Engineering, Dankook University, Gyeonggi, Republic of Korea

Corresponding Author:
Han Gil Seo, MD, PhD
Department of Rehabilitation Medicine
Seoul National University Hospital
101 Daehak-Ro Jongno-Gu
Seoul, 03080
Republic of Korea
Phone: 82 220721659
Fax: 82 260725244
Email: hgseo80@gmail.com

Abstract

Background: Cognitive training using virtual reality (VR) may result in motivational and playful training for patients with mild cognitive impairment and mild dementia. Fully immersive VR sets patients free from external interference and thus encourages patients with cognitive impairment to maintain selective attention. The enriched environment, which refers to a rich and stimulating environment, has a positive effect on cognitive function and mood.

Objective: The aim of this study was to investigate the feasibility and usability of cognitive training using fully immersive VR programs in enriched environments with physiatrists, occupational therapists (OTs), and patients with mild cognitive impairment and mild dementia.

Methods: The VR interface system consisted of a commercialized head-mounted display and a custom-made hand motion tracking module. We developed the virtual harvest and cook programs in enriched environments representing rural scenery. Physiatrists, OTs, and patients with mild cognitive impairment and mild dementia received 30 minutes of VR training to evaluate the feasibility and usability of the test for cognitive training. At the end of the test, the usability and feasibility were assessed by a self-report questionnaire based on a 7-point Likert-type scale. Response time and finger tapping were measured in patients before and after the test.

Results: Participants included 10 physiatrists, 6 OTs, and 11 patients with mild cognitive impairment and mild dementia. The mean scores for overall satisfaction with the program were 5.75 (SD 1.00) for rehabilitation specialists and 5.64 (SD 1.43) for patients. The response time of the dominant hand in patients decreased after the single session of cognitive training using VR, but this was not statistically significant ($P=.25$). There was no significant change in finger tapping in either the right or left hand ($P=.48$ and $P=.42$, respectively). None of the participants reported headaches, dizziness, or any other motion sickness after the test.

Conclusions: A fully immersive VR cognitive training program may be feasible and usable in patients with mild cognitive impairment and mild dementia based on the positive satisfaction and willingness to use the program reported by physiatrists, OTs, and patients. Although not statistically significant, decreased response time without a change in finger tapping rate may reflect a temporary increase in attention after the test. Additional clinical trials are needed to investigate the effect on cognitive function, mood, and physical outcomes.

Introduction

Dementia, a major global public health burden, is a neurodegenerative disorder that impairs cognitive functions such as memory, language, and goal-directed behaviors [1]. Cognitive training in early stages of dementia has been considered a promising tool for addressing the impact of cognitive changes [2]. A recent systematic review demonstrated that computerized cognitive training improves cognition and psychosocial function in patients with mild cognitive impairment; however, there is lack of evidence in those with dementia [3]. Patients with cognitive impairment are less likely to engage and less able to concentrate in repetitive training compared with the healthy elderly population [4]. On the other hand, cognitive training based on virtual reality (VR) introduces motivational and playful aspects of training for patients with mild cognitive impairment and dementia [5].

Emerging VR cognitive training focuses on orientation, spatial navigation, face recognition, memory function, and instrumental activities of daily living [6]. The advantage of applying VR is that activities, tasks, and evaluation can occur in a secure environment [7]. Furthermore, VR training parameters can be adjusted within a person-centered therapeutic milieu [8]. A multisensorial experience through VR enables feedback-based learning [9].

Although virtual environments are artificial in nature, the user can elicit a rich feeling of realness and agency [10]. The immersion level of the VR system has a critical role on visual realism. The sense of "presence" refers to being and therefore how well the virtual environment represents the real world [11]. Reality level refers to the level at which a user truly experiences immersion [12]. Fully immersive VR sets people free from external interference. This VR world encourages patients with cognitive impairment to maintain selective attention [13]. However, VR technologies with sufficient levels of immersion or interaction have not been developed for this patient population [6].

Familiar image-based virtual environments can stimulate the recollections of autobiographical memory in healthy elderly [14]. Enriched environments provide conditions that enhance sensory, cognitive, and motor stimulation [15]. Furthermore, experiencing rich environments evokes positive emotions. A positive feeling is a protective factor against brain dysfunction, similar to that of the cognitive enrichment hypothesis [16]. It has been reported that enriched environments improve spatial impairment and memory deficits in vascular dementia and Alzheimer disease mouse models [17,18]. Recently, a video game that applied enriched environments was found to have a positive effect on cognitive function and mood in subjects with normal aging and mild to moderate cognitive impairment [4].

Only a few studies on cognitive training using fully immersive VR for patients with mild cognitive impairment and dementia have been conducted [13,19]. VR games based on enriched environments can be helpful for full immersion and comfort of patients experiencing cognitive decline. Furthermore, as it targets the elderly with cognitive impairment, it is necessary to develop VR games that reflect cultural specificity. We developed a fully immersive VR system based on enriched environments in order to train attention, memory, and executive function in the elderly. The purpose of this study was to investigate the feasibility and usability of cognitive training using the VR system with rehabilitation specialists and patients with mild cognitive impairment and mild dementia.

Methods

Study Design

This was a pilot study to evaluate the usability and feasibility of a fully immersive VR cognitive training program with an enriched environment. The study consisted of a single session and a survey completed by physiatrists and occupational therapists (OTs) as well as patients with mild cognitive impairment and mild dementia. This study protocol was approved by the Institutional Review Board of Seoul National University Hospital (IRB No. 1809-126-975) on October 16, 2018. All participants provided written informed consent. The study was performed under the principles of Good Clinical Practice and the Helsinki Declaration.

VR Interface System

The VR interface system consisted of an HTC Vive head-mounted display and a custom-made hand motion tracking module developed for hand pose estimation and 3-dimensional positions of the hands in the working space.

Figure 1 shows the overall system architecture of the module for hand pose estimation in a virtual environment. The hand motion tracking module has a camera on the palm side of the hand to capture images; the camera is synced with a computer system for the deep learning process in order to estimate the hand poses. Each of the 2 modules for both hands transmits real-time hand images to the deep learning process in the computer system through WiFi communication. The deep learning process computes finger joint positions from the images to estimate hand poses in real time, which is delivered to the interface synchronization process in the computer system for the VR game through Bluetooth. HTC Vive trackers are installed on the backside of the hands to obtain the hands' positions, which are also used to control the VR game.

Figure 1. The overall system architecture of the module for hand pose estimation: (A) motion tracking module, (B) camera module, (C) transmission/reception module and battery.
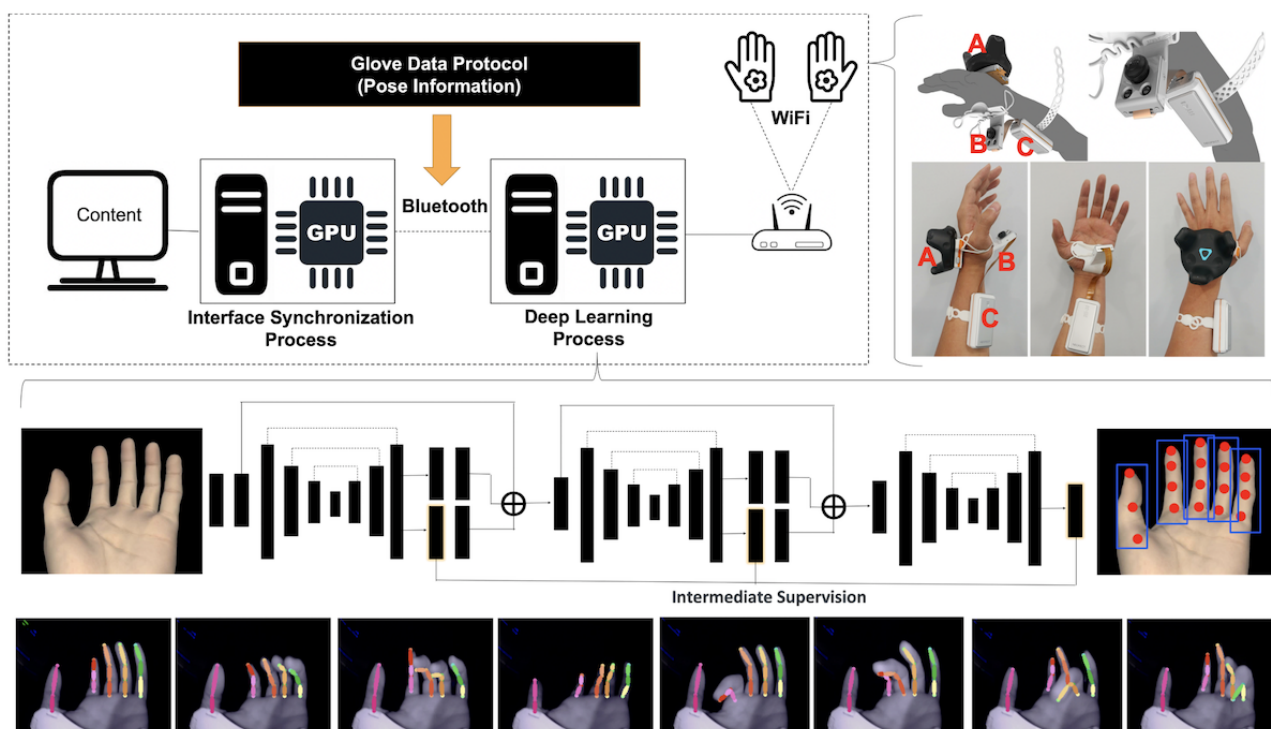
The network model of our deep learning process was designed based on the hourglass network [20]. The hourglass network has primarily been used for human body pose estimation. Therefore, the structure of the network is not optimized for hand pose estimation. We mainly focused on the modification of the feature extraction layers of the network so that finger features could be efficiently extracted. However, as it was difficult to obtain sufficient hand image data for training, we used the composite hand data generator to generate simulated hand image data [21]. More than 2 million images were generated to train our deep learning process. As shown at the bottom of Figure 1, the 2-dimensional joint position of each finger was successfully extracted using our proposed method.

With the joint position data, we could estimate the finger joint angles from an inverse kinematics model. Moreover, without the inverse kinematics model, gestures like grip could be easily recognized from the joint position data. For example, grip could be detected if the vertical positions of fingertips were lower than the threshold value. Thumb-to-finger tap gesture could also be easily detected using only the 2-dimensional joint position data [22]. These data were then used to interface with the VR game.

The hand motion tracking module was designed such that patients could wear it easily without much effort. The silicon pad and strap made the module easy to wear, remove, and clean.

VR Cognitive Training Program Based on an Enriched Environment

We developed virtual harvest and cook games in enriched environments representing Korean rural scenery. In all the games, the patients interacted with the environment from an egocentric point of view (known as "first person point of view"). The developed algorithm is shown in Multimedia Appendix 1.

In the default scene of the program, the user was standing in a stream with a view of a country field and house (Figure 2). Sounds of water, wind, birds, and soft music were played as background music. The scene mimicked an environment familiar and comfortable for Korean elderly patients.

The harvest game aimed to improve sustained, alternating, and selective attention and working memory (Figure 3). In the game, there were 4 backgrounds: farm A (chilis, tomatoes, and cucumbers), farm B (strawberries, paprika, and eggplants), an orchard (apples, mandarins, and pears), and a hen house (brown eggs and white eggs). The type and number of crops and eggs harvested and the time limit were controlled autonomously by the operator, depending on each participant's performance level. After the operator inputted the harvest target and time frame, an artificial voice informed the patient. If the participants were unable to complete the task in the appointed time, visual cues with green lights around the object were provided. If the participant performed incorrectly, there was no limit but it remained on record. After the participants finished the task, the instructor checked time taken by the participant to complete the task, the number of times the direction was replayed, and the harvest type and number.

The cook game focused on improvements in sustained and selective attention; working, spatial, and procedure memory; and executive function. The cook program included 3 types of recipes: fried eggs, gimbap (dried seaweed roll), and soybean paste stew. The fried egg recipe consisted of 12 steps, gimbap recipe consisted of 11 steps, and soybean paste stew recipe consisted of 14 steps. The recipe was selected by the operator depending on the participant's level. The participants were verbally instructed by the program on how to execute the allocated recipe. If a participant could not cook the recipe properly, a green arrow around the item necessary for the present

step glowed. The strategies of the cook game were based on "errorless learning" [23]. If patients could not progress to the next step, the program did not progress to the next step. After the participants performed the task, the execution time was tabulated for each step.

The VR cognitive training program was evaluated and improved through the initial usability test by 12 physiatrists and 7 OTs before the final application.

Figure 2. Default scenes in the program: (A) primary scene showing a usual Korean farm village and country house where one could harvest and cook and (B) patient information input scene.

XSL•FO
RenderX

Figure 3. Harvest and cook game: (A) farm A (chilis, tomatoes, and cucumbers), (B) farm B (strawberries, paprika, and eggplants), (C) orchard (apples, mandarins, and pears), (D) hen house (brown eggs and white eggs), (E) fried eggs, (F) gimbap (dried seaweed roll), (G) soybean paste stew, (H) results of the cooking game.

Procedure

Participants including physiatrists, OTs, and patients were provided with a description of the program objectives and procedures before the test. All participants were seated on a fixed chair, wearing hand motion tracking modules, and supervised by an OT during the test (Multimedia Appendix 2). The therapist was trained to manipulate the software for the usability test and to deal with adverse events. The OT adjusted game difficulty to match the patient's level based on the Mini-Mental State Examination (MMSE) score and Clinical Dementia Rating (CDR). Physical and verbal assistance was provided for participants who struggled to understand the instructions properly. The participants experienced all 4 harvest games and 3 cooking games. The usability and feasibility test was a single session lasting of 30 minutes.

Participants

Physiatrists and OTs with more than 1 year of clinical experience were recruited to test the usability and feasibility of the developed program. Patients were recruited from a tertiary hospital. A total of 15 patients were screened for eligibility for study inclusion. Inclusion criteria were age ≥65 years and a score of more than 19 and less than 28 points on the MMSE. Exclusion criteria were patients with moderate to severe dementia who could not understand the research content or agree voluntarily, a history of severe dizziness or epilepsy, psychiatric symptoms or behavioral problems that made it difficult to participate in the study, and other severe medical problems such as neurological or orthopedic diseases.

Outcome Measures

Physiatrists and OTs evaluated the usability of the developed VR cognitive training system using a self-report questionnaire based on a 7-point Likert-type scale. Questionnaire items consisted of 6 categories and 22 items: overall satisfaction, acceptability (needs, favorability, effectualness for mild cognitive impairment and mild dementia, and distinction), satisfactoriness (intent adequacy, setting adequacy, setting immediacy, inducement of interest, and content diversity), expectation effectiveness (attraction of voluntary participation from patients, labor reduction for therapists, and contribution to therapists' work), stability (confirmation of proper equipment positioning, warning of system errors, external appearance of safety, durability, and stability satisfaction), and others (willingness to use, additional use intention with other therapeutic tools, and recommendation intention). The higher the score, the more positive the result. The survey also included additional suggestions regarding the developed system.

The patients' baseline characteristics, including age, sex, medical history, MMSE, CDR, and Geriatric Depression Scale (GDS), were assessed. At the end of the experience, the feasibility of the enriched environment VR games was assessed by patients using a self-report questionnaire with a 7-point Likert-type scale based on 9 items (overall satisfaction, interest, mood, motivation, difficulty, discomfort, anxiety, willingness to use, and expectations for VR rehabilitation). The survey also included additional descriptions of subjective experiences.

Reaction time was assessed to evaluate the activity process of the central nervous system, such as motor preparation and motor program [24]. Using an iPad app (Reaction Test & Speed Test, Wang Haiwen), the instructor instructed patients that when the red screen became green, the screen should be touched as quickly as possible. This app measured response time from the color change to the screen touch. Patients tried for a total of 5 times using their dominant hand.

To assess motor speed, the finger tapping test was used [25]. Participants tapped just an iPad (digital finger tapping test, SyBu Data) with the index finger as fast as they could for 10 seconds. An average of 3 trials with both hands was completed.

Statistical Analysis

Baseline characteristics and assessment data are expressed as mean and SD for continuous variables. The Wilcoxon signed-ranks test was used to evaluate changes in results before and after the usability test. The Mann-Whitney test was employed to compare the number of finger taps between the right and left hands. A P value <.05 was considered statistically significant. SPSS Statistics 21.0 for Windows (IBM Corp, Armonk, NY) was used for all analyses.

Results

In this study, 10 physiatrists and 6 OTs (mean experience: 4.9 years, range 2-30 years) participated on March 21, 2019. The overall satisfaction score of the system was 5.75 (SD 1.00). In descending order, the mean score was highest for acceptability (5.79, SD 0.16), expectation effectiveness (5.32, SD 0.43), satisfactoriness (5.11, SD 0.79), and stability (4.86. SD 0.32; Table 1).

A total of 11 patients with mild cognitive impairment and mild dementia were enrolled in this study between April 26, 2019 and May 17, 2019. The average age of the participants was 72.64 years (SD 4.65 years; Table 2). The MMSE score ranged from 23 to 28 points, and the mean MMSE score was 26.91 points (SD 1.58 points). The CDR was 0.5 in 8 patients and stage 1 in 3 patients. The mean baseline GDS score was 17.55 points (SD 6.30 points, range 10-30 points). All patients were right-handed.

XSL·FO

RenderX

Table 1. Physiatrists' and occupational therapists' self-report questionnaire results on a 7-point Likert-type scale. The higher the score, the more positive the result.

Scale and subscale	Score, mean (SD)
Overall satisfaction	5.75 (1.00)
Acceptability	5.79 (0.16)
Needs	5.88 (0.89)
Favorability	6.00 (0.82)
Effectualness (mild cognitive impairment)	5.63 (0.89)
Effectualness (mild dementia)	5.63 (0.81)
Distinction	5.81 (1.17)
Satisfactoriness	5.11 (0.79)
Intent adequacy	5.38 (0.89)
Setting adequacy	4.19 (1.33)
Setting immediacy	4.50 (1.37)
Inducement of interest	6.19 (0.75)
Content diversity	5.31 (1.54)
Expectation effectiveness	5.32 (0.43)
Attraction of voluntary participation from participants	5.75 (1.06)
Labor reduction for therapists	4.89 (1.61)
Contribution to therapists' work	5.31 (0.95)
Stability	4.86 (0.32)
Confirmation of proper equipment positioning	5.00 (1.46)
Warning of system errors	4.50 (1.60)
External appearance of safety	5.00 (1.10)
Durability	4.56 (1.26)
Stability satisfaction	5.25 (1.24)
Others	5.79 (0.14)
Willingness to use	5.63 (1.41)
Additional use intention with other therapeutic tools	5.88 (0.89)
Recommendation intention	5.88 (0.81)

Table 2. Patients' demographics and baseline characteristics (n=11).

Number	Age (years)	Sex	MMSE[a] (points)	CDR[b] (points)	GDS[c] (points)
1	72	M[d]	28	0.5	22
2	74	F[e]	28	0.5	30
3	71	F	27	1	21
4	76	M	27	0.5	14
5	81	M	25	0.5	16
6	76	M	28	1	24
7	67	F	28	0.5	10
8	76	F	28	0.5	15
9	73	M	23	1	10
10	67	F	27	0.5	12
11	66	F	27	0.5	19

[a]MMSE: Mini-Mental State Examination.

[b]CDR: Clinical Dementia Rating.

[c]GDS: Geriatric Depression Scale.

[d]M: male.

[e]F: female.

For overall satisfaction with the program, the mean patient rating was 5.64 (SD 1.43). The study items with high scores were comfort (6.91, SD 0.30), anxiety (6.27, SD 1.62), and mood (6.18, SD 1.40). The remainder of the scores were as follows: interest, 5.82 (SD 1.33); motivation, 5.36 (SD 1.57); difficulty, 5.45 (SD 1.51); willingness to use, 5.82 (SD 1.83); and expectations for VR rehabilitation, 5.55 (SD 1.44). The mean response time of the dominant hand decreased after a single session of cognitive training using VR, but the difference was not statistically significant (pre-VR: 612.18 ms, SD 186.35 ms; post-VR: 546.82 ms, SD 130.64 ms; P=.25). There was no change in finger taps of the right (pre-VR: 37.97, SD 11.93; post-VR: 38.59, SD 11.75; P=.48) or left hand (pre-VR: 35.81, SD 11.59; post-VR: 33.85, SD 10.82; P=.42). There was no significant difference in the mean number of finger taps between the right and left hands before (right: 37.97, SD 11.93; left: 35.81, SD 11.59; P=.56) and after (right: 38.59, SD 11.75; left: 33.85, SD 10.82; P=.30) the training.

None of the participants reported headaches, dizziness, or any other form of motion sickness after the test.

Discussion

Principal Findings

We developed a memory and attention training program using enriched VR environments for patients with mild cognitive impairment and mild dementia. Our results suggest that the VR cognitive training system is feasible and usable based on the feedback received from physiatrists, OTs, and patients with mild cognitive impairment and mild dementia. None of the participants complained of motion sickness, and no adverse events occurred. Although not statistically significant, the decreased response time without changing the rate of finger

tapping in patients may reflect a temporary increase in attention after the test.

Fully immersive VR can induce psychological immersion and high presence. Lifelike VR may narrow the gap between reality and the virtual world [26]. The high level of immersion and visual realism trigger autobiographical memories [27]. A combination of the enriched environment, which is almost like reality, and VR may enable patients with cognitive decline to feel emotionally stable. Neuropsychiatric symptoms including depression, apathy, and agitation occur in most patients with dementia [28]. The patients in this study were found to be mildly depressive as per their GDS score. After the test, the patients gave positive feedback on the questionnaire with regards to comfort, anxiety, and mood. Therefore, the program might be helpful due to improvement in mood in patients with mild cognitive impairment and mild dementia. This finding is in line with a report stating psychosocial intervention through individualized reminiscence and multisensory stimulation evoked emotional and social benefits for those with dementia, along with a preserved sense of identity [29].

In this study, allowing free hand movements in elderly patients who are not used to the VR machine operation ensured that they could access the VR environment more comfortably. Physiatrists, OTs, and patients did not report incommodiousness of a custom-made hand module. However, 3 rehabilitation specialists suggested that it would be better to add haptic feedback when grasping objects and to enhance the sensitivity of the motion detector. To bring about interactions between the user and the virtual environment, several input devices are needed [30]. A potential problem with hand-held devices is that a stiff controller limits kinematic error and completes the actions without continuous participation from the patient, which may limit the effect of training. Furthermore, an unsolved problem regarding the motion tracking systems and instrumented gloves

is that they are limited when creating accurate user movements in the virtual environment [31]. In addition, bacterial contamination and sterilization could be an issue for medical use of hand-held devices and instrumented gloves. There is an ongoing need to make intuitive input devices for the freedom to manipulate devices and increase accuracy. Although the hand motion tracking module we developed is a prototype, it could compensate for several shortcomings of existing products.

It is also necessary to consider the needs of caregiving staff when designing and developing interventional VR programs. Physiatrists and rehabilitation therapists have many training interface options to choose from, including VR, video gaming, and other tablet-based applications [32]. A survey of rehabilitation specialists showed high overall satisfaction and intent to use the VR system for further training. However, they reported that the hardware system needs to be improved with regards to setting the speed and adequacy. It is also worth considering incorporating a haptic device into the hardware system to make patients with cognitive impairment more comfortable with their VR tasks and to provide them the appropriate tactile stimuli. Because the hardware system is not commercialized, continuous development is needed to improve user convenience.

Limitations

This study has several limitations. First, the short test time may be insufficient to influence cognitive components and motor function. Because the primary purpose of this study was to evaluate the usability and feasibility of the developed system, all participants underwent a single session of VR cognitive training. Response time decreased after the short-term VR experience in patients although it was not statistically significant. This result may suggest improved cognitive process without changes in motor speed after VR cognitive training, as no changes were noted in the finger tapping test. To evaluate the efficacy of the developed program, a further clinical trial should be designed including several training sessions and a conventional cognitive training control group. Second, motion sickness was not evaluated objectively in this study. In order to minimize motion sickness during the VR experience, the researchers observed the condition of participants carefully and continuously asked them about any inconvenience. Participants in this study did not complain of motion sickness. Nevertheless, cybersickness due to VR use must be considered, especially when fully immersive VR is being utilized in cognitively impaired patients [33]. An objective assessment that can measure motion sickness symptoms such as dizziness, nausea, and headache may be required in future studies. Last, the developed program was only used with those with mild cognitive impairment and mild dementia. This study population was chosen based on previous studies that have reported the effect of computerized cognitive training in people with cognitive impairment [3]. However, as the enriched VR environment may be beneficial for reminiscence, relaxation, and enjoyment for different stages of dementia, future studies are warranted in patients with moderate and severe dementia.

Conclusions

We developed a fully immersive VR cognitive training program with an enriched environment for patients with mild cognitive impairment and mild dementia. The feasibility and usability of the program were verified based on the positive satisfaction and willingness to use reported by physiatrists, OTs, and patients. Although we were not able to identify functional changes with the use of a single session, there may be an increase in attention following training. Additional clinical trials are needed to confirm the effects of this program on cognitive function, mood, and physical outcomes.

Conflicts of Interest

DY and YC are employees of Neofect (Gyeonggi-do, Korea), which made the VR interface system in this study.

Multimedia Appendix 1
Algorithm for program development.
[PNG File , 181 KB - games_v8i4e18127_app1.png]

Multimedia Appendix 2
A patient experiences the virtual reality training with the head mount display (blue circle) and hand motion tracking module (red circles).
[PNG File , 1987 KB - games_v8i4e18127_app2.png]

References

1. World Health Organization. Dementia: key facts. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2020-09-23]

2. Rodakowski J, Saghafi E, Butters MA, Skidmore ER. Non-pharmacological interventions for adults with mild cognitive impairment and early stage dementia: An updated scoping review. Mol Aspects Med 2015 Jun;43-44:38-53 [FREE Full text] [doi: 10.1016/j.mam.2015.06.003] [Medline: 26070444]

3. Hill NTM, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A. Computerized Cognitive Training in Older Adults With Mild Cognitive Impairment or Dementia: A Systematic Review and Meta-Analysis. Am J Psychiatry 2017 Apr 01;174(4):329-340. [doi: 10.1176/appi.ajp.2016.16030360] [Medline: 27838936]

4. Ben-Sadoun G, Sacco G, Manera V, Bourgeois J, König A, Foulon P, et al. Physical and Cognitive Stimulation Using an Exergame in Subjects with Normal Aging, Mild and Moderate Cognitive Impairment. J Alzheimers Dis 2016 Jun 30;53(4):1299-1314. [doi: 10.3233/JAD-160268] [Medline: 27372645]

5. McEwen D, Taillon-Hobson A, Bilodeau M, Sveistrup H, Finestone H. Two-week virtual reality training for dementia: Single case feasibility study. J Rehabil Res Dev 2014;51(7):1069-1076 [FREE Full text] [doi: 10.1682/JRRD.2013.10.0231] [Medline: 25437527]

6. García-Betances RI, Arredondo Waldmeyer MT, Fico G, Cabrera-Umpiérrez MF. A succinct overview of virtual reality technology use in Alzheimer's disease. Front Aging Neurosci 2015;7:80 [FREE Full text] [doi: 10.3389/fnagi.2015.00080] [Medline: 26029101]

7. Subramanian SK, Lourenço CB, Chilingaryan G, Sveistrup H, Levin MF. Arm motor recovery using a virtual reality intervention in chronic stroke: randomized control trial. Neurorehabil Neural Repair 2013 Jan;27(1):13-23. [doi: 10.1177/1545968312449695] [Medline: 22785001]

8. Rose V, Stewart I, Jenkins KG, Tabbaa L, Ang CS, Matsangidou M. Bringing the outside in: The feasibility of virtual reality with people with dementia in an inpatient psychiatric care setting. Dementia (London) 2019 Sep 11:1471301219868036. [doi: 10.1177/1471301219868036] [Medline: 31510801]

9. Schultheis MT, Rizzo AA. The application of virtual reality technology in rehabilitation. Rehabil Psychol 2001 Mar 21;46(3):296-311. [doi: 10.1037/0090-5550.46.3.296]

10. Riva G, Castelnuovo G, Mantovani F. Transformation of flow in rehabilitation: the role of advanced communication technologies. Behav Res Methods 2006 May;38(2):237-244. [doi: 10.3758/bf03192775] [Medline: 16956100]

11. Nash EB, Edwards GW, Thompson JA, Barfield W. A Review of Presence and Performance in Virtual Environments. International Journal of Human-Computer Interaction 2000 May;12(1):1-41. [doi: 10.1207/s15327590ijhc1201_1]

12. Baños R, Botella C, Garcia-Palacios A, Villa H, Perpiña C, Alcañiz M. Presence and Reality Judgment in Virtual Environments: A Unitary Construct? 2000 Jun;3(3):327-335. [doi: 10.1089/10949310050078760]

13. Optale G, Urgesi C, Busato V, Marin S, Piron L, Priftis K, et al. Controlling memory impairment in elderly adults using virtual reality memory training: a randomized controlled pilot study. Neurorehabil Neural Repair 2010 May;24(4):348-357. [doi: 10.1177/1545968309353328] [Medline: 19934445]

14. Benoit M, Guerchouche R, Petit P, Chapoulie E, Manera V, Chaurasia G, et al. Is it possible to use highly realistic virtual reality in the elderly? A feasibility study with image-based rendering. Neuropsychiatr Dis Treat 2015;11:557-563 [FREE Full text] [doi: 10.2147/NDT.S73179] [Medline: 25834437]

15. Nithianantharajah J, Hannan AJ. Enriched environments, experience-dependent plasticity and disorders of the nervous system. Nat Rev Neurosci 2006 Sep;7(9):697-709. [doi: 10.1038/nrn1970] [Medline: 16924259]

16. Hertzog C, Kramer AF, Wilson RS, Lindenberger U. Enrichment Effects on Adult Cognitive Development: Can the Functional Capacity of Older Adults Be Preserved and Enhanced? Psychol Sci Public Interest 2008 Oct;9(1):1-65. [doi: 10.1111/j.1539-6053.2009.01034.x] [Medline: 26162004]

17. Jin X, Li T, Zhang L, Ma J, Yu L, Li C, et al. Environmental Enrichment Improves Spatial Learning and Memory in Vascular Dementia Rats with Activation of Wnt/β-Catenin Signal Pathway. Med Sci Monit 2017 Jan 13;23:207-215 [FREE Full text] [doi: 10.12659/msm.902728] [Medline: 28082734]

18. Balthazar J, Schöwe NM, Cipolli GC, Buck HS, Viel TA. Enriched Environment Significantly Reduced Senile Plaques in a Transgenic Mice Model of Alzheimer's Disease, Improving Memory. Front Aging Neurosci 2018;10:288 [FREE Full text] [doi: 10.3389/fnagi.2018.00288] [Medline: 30319394]

19. Manera V, Chapoulie E, Bourgeois J, Guerchouche R, David R, Ondrej J, et al. A Feasibility Study with Image-Based Rendered Virtual Reality in Patients with Mild Cognitive Impairment and Dementia. PLoS One 2016 Mar 18;11(3):e0151487 [FREE Full text] [doi: 10.1371/journal.pone.0151487] [Medline: 26990298]

20. Newell A, Yang K, Deng J. Stacked Hourglass Networks for Human Pose Estimation. 2016 Presented at: European Conference on Computer Vision; October 8-16, 2016; Amsterdam, The Netherlands. [doi: 10.1007/978-3-319-46484-8_29]

21. Yang D, Moon BS, Kim H, Choi Y. Synthetic Hands Generator for RGB Hand Tracking. 2018 Presented at: TENCON 2018; October 28-31, 2018; Jeju, South Korea. [doi: 10.1109/tencon.2018.8650519]

22. Yang D, Choi Y. Palm glove: wearable glove based on palm-camera for thumb-to-finger tap recognition. 2018 Presented at: International Conference on Information and Communication Technology Convergence (ICTC); October 17-19, 2018; Jeju, South Korea. [doi: 10.1109/ictc.2018.8539616]

23. Foloppe DA, Richard P, Yamaguchi T, Etcharry-Bouyx F, Allain P. The potential of virtual reality-based training to enhance the functional autonomy of Alzheimer's disease patients in cooking activities: A single case study. Neuropsychol Rehabil 2018 Jul;28(5):709-733. [doi: 10.1080/09602011.2015.1094394] [Medline: 26480838]

XSL•FO
RenderX

24. Pourazar M, Mirakhori F, Hemayattalab R, Bagherzadeh F. Use of virtual reality intervention to improve reaction time in children with cerebral palsy: A randomized controlled trial. Dev Neurorehabil 2018 Nov;21(8):515-520. [doi: 10.1080/17518423.2017.1368730] [Medline: 28933977]

25. Schmitt L. Finger-Tapping Test. In: Volkmar FR, editor. Encyclopedia of Autism Spectrum Disorders. New York, NY: Springer Publishing Company; 2013:1296.

26. Yeh SC, Chen YC, Tsai CF, Rizzo A. An innovative virtual reality system for mild cognitive impairment: Diagnosis and evaluation. 2012 Presented at: IEEE-EMBS Conference on Biomedical Engineering and Sciences; December 17-19, 2012; Langkawi, Malaysia. [doi: 10.1109/iecbes.2012.6498023]

27. D'Cunha NM, Nguyen D, Naumovski N, McKune AJ, Kellett J, Georgousopoulou EN, et al. A Mini-Review of Virtual Reality-Based Interventions to Promote Well-Being for People Living with Dementia and Mild Cognitive Impairment. Gerontology 2019;65(4):430-440 [FREE Full text] [doi: 10.1159/000500040] [Medline: 31108489]

28. Lyketsos CG, Lopez O, Jones B, Fitzpatrick AL, Breitner J, DeKosky S. Prevalence of neuropsychiatric symptoms in dementia and mild cognitive impairment: results from the cardiovascular health study. JAMA 2002 Sep 25;288(12):1475-1483. [doi: 10.1001/jama.288.12.1475] [Medline: 12243634]

29. Goodall G, Ciobanu I, Broekx R, Sørgaard J, Anghelache I, Anghelache-Tutulan C, et al. Using Adaptive Immersive Environments to Stimulate Emotional Expression and Connection in Dementia Care. 2018 Presented at: Fourth International Conference on Human and Social Analytics; June 24-28, 2018; Venice, Italy.

30. Adamovich SV, Fluet GG, Tunik E, Merians AS. Sensorimotor training in virtual reality: a review. NeuroRehabilitation 2009;25(1):29-44 [FREE Full text] [doi: 10.3233/NRE-2009-0497] [Medline: 19713617]

31. Wolbrecht ET, Chan V, Reinkensmeyer DJ, Bobrow JE. Optimizing compliant, model-based robotic assistance to promote neurorehabilitation. IEEE Trans Neural Syst Rehabil Eng 2008 Jun;16(3):286-297. [doi: 10.1109/TNSRE.2008.918389] [Medline: 18586608]

32. Levac D, Dumas HM, Meleis W. A Tablet-Based Interactive Movement Tool for Pediatric Rehabilitation: Development and Preliminary Usability Evaluation. JMIR Rehabil Assist Technol 2018 Nov 26;5(2):e10307 [FREE Full text] [doi: 10.2196/10307] [Medline: 30478025]

33. Nichols S, Patel H. Health and safety implications of virtual reality: a review of empirical evidence. Appl Ergon 2002 May;33(3):251-271. [doi: 10.1016/s0003-6870(02)00020-0] [Medline: 12164509]

Abbreviations

CDR: Clinical Dementia Rating
F: female
GDS: Geriatric Depression Scale
M: male
MMSE: Mini-Mental State Examination
OT: occupational therapist
VR: virtual reality

XSL•FO
RenderX

Viewpoint

A Primer on Usability Assessment Approaches for Health-Related Applications of Virtual Reality

Timothy Zhang[1,2]; Richard Booth[3], RN, PhD; Royce Jean-Louis[4], MHI; Ryan Chan[3], RN, MScN; Anthony Yeung[5], MD; David Gratzer[1,5], MD; Gillian Strudwick[1,4], RN, PhD

[1]Centre for Addiction and Mental Health, Toronto, ON, Canada

[2]Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[3]Arthur Labatt Family School of Nursing, Western University, London, ON, Canada

[4]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[5]Department of Psychiatry, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Gillian Strudwick, RN, PhD
Centre for Addiction and Mental Health
Bell Gateway Building, Room 7343
1001 Queen St W
Toronto, ON, M6J 1H4
Canada
Phone: 1 4165358501 ext 39333
Email: gillian.strudwick@camh.ca

Abstract

Health-related virtual reality (VR) applications for patient treatment, rehabilitation, and medical professional training are on the rise. However, there is little guidance on how to select and perform usability evaluations for VR health interventions compared to the supports that exist for other digital health technologies. The purpose of this viewpoint paper is to present an introductory summary of various usability testing approaches or methods that can be used for VR applications. Along with an overview of each, a list of resources is provided for readers to obtain additionally relevant information. Six categories of VR usability evaluations are described using a previously developed classification taxonomy specific to VR environments: (1) cognitive or task walkthrough, (2) graphical evaluation, (3) post hoc questionnaires or interviews, (4) physical performance evaluation, (5) user interface evaluation, and (6) heuristic evaluation. Given the growth of VR in health care, rigorous evaluation and usability testing is crucial in the development and implementation of novel VR interventions. The approaches outlined in this paper provide a starting point for conducting usability assessments for health-related VR applications; however, there is a need to also move beyond these to adopt those from the gaming industry, where assessments for both usability and user experience are routinely conducted.

KEYWORDS

Introduction

In the last decade, there has been a tremendous increase in the use of virtual reality (VR) technology in a variety of global contexts, including entertainment (eg, gaming), education, marketing, and design. VR broadly describes digitally created simulations where a person can be immersed in a computer-generated reality and complete tasks or interact with a virtual environment. Equipment such as VR headsets that allow individuals to experience the sounds and sights of a virtual world are often utilized to create an immersive experience.

More recently, numerous applications of VR specific to the health context have been identified and used [1-7], as research involving VR for health-related applications is gaining interest. As of July 2020, over 1000 studies were registered on ClinicalTrials.gov—a registry of clinical trials in the United States—for assessing VR interventions, such as anxiety management, distraction during painful procedures, gait training, rehabilitation, phobias, and medical education [8]. VR has been shown to be able to act as a low-cost and effective analgesic for pain arising in cases such as invasive medical procedures or even cancer in pediatric patients [9-11]. Hospitals may be able to leverage VR to reduce preoperative anxiety in patients, as

well as a treatment method for those with generalized anxiety disorder [12]. A recent study by Donker et al showed that patients with acrophobia who received exposure therapy through a gamified, VR-enabled, self-help app had significant reductions in acrophobic symptoms [13]. Notably, in addition to the lack of need for a psychiatrist to be directly present during this intervention, the total cost per patient came to approximately US $24 through the use of Google Cardboard as the VR headset [13], exemplifying the ability of VR to increase treatment access while also significantly reducing costs. These examples only scratch the surface of the exciting potential of VR in health care.

To complement the significant amount of benefits that VR applications bring to health-related contexts, a focus on the usability of health information technologies needs to be maintained, particularly given the diverse needs and abilities of the user base (eg, patients, health professionals, family members, etc). By usability, we refer to how easily the technology can be utilized by an individual based on three cycles or steps [14]. Often, the effort invested into ensuring the usability of a technology or application goes unnoticed until the user interacts with a poorly designed system. A user's proficiency with a technology may originate from a combination of their own self-exploratory learning as well as more formal, structured lessons and walkthroughs. Given the novel nature of VR in health care, the likely paucity of the latter places a greater emphasis on ensuring VR technologies are intuitive and easy to adopt for those who are new to the technology.

In the context of VR specifically, this includes both the use of the hardware (eg, headset) as well as the immersive software and VR experience as perceived by the user.

Purpose

The purpose of this viewpoint paper is to conduct the following: (1) highlight the need to conduct usability assessments for VR apps, (2) provide a primer on the potential usability assessment approaches that can be applied to VR in health-related contexts and their potential challenges, and (3) direct readers to several resources where additional information on the topic can be found.

The Need to Conduct Usability Assessments for VR Health-Related Applications

One of the challenges of VR for health-related applications is assessing and addressing issues related to usability. Health-related applications of VR may warrant an even greater focus on usability testing than nonhealth-related applications, given that the user base (ie, those typically with illnesses, chronic conditions, or disabilities) is diverse in terms of ages, abilities, and beyond, and may have special needs that need to be accounted for when utilizing the technology. In addition, one of the most common problems associated with VR is motion sickness, which is often related to the quality of the virtual space mapping to the replicated physical setting [15]. This can be a significant barrier to users looking to obtain health-related benefits from using VR. Yet, there are methods in which motion

sickness may be evaluated and addressed before the technology is implemented. The integration of VR into treatment plans can also meet commonly seen elements of friction associated with new technologies, such as distrust during adoption, although in some situations these can dwindle following introductory exposure [16]. Other limitations of contemporary VR include the challenge of generating varied types of tactile sensations [17] and other types of multisensory integration [18].

Assessment approaches for analyzing and evaluating the usability of various VR technologies for health-related applications have generally been understudied and not well described in the research literature. We conducted a cursory search ourselves of several academic databases and found limited explanations of usability methods utilized in the development stage of VR applications and an even more limited body of literature on how to conduct usability assessments for VR used for health-related purposes. While reasons for this gap in knowledge are likely due to the nascent nature of the field, further work must be completed toward generating best practices related to VR usability to assist practitioners and researchers in the development and diffusion of these sorts of innovations. For instance, outside of the VR context, there is an extensive literature base identifying the need for technologies that are used for health-related applications to be *user friendly* and have a high degree of *ease of use*, often incorporating in lessons from the human factors discipline [19-22]. Numerous papers, including one reporting on the System Usability Scale [23], have been published describing ways to assess usability for non-VR technologies, including electronic health records and mobile health apps [24-26]. Yet, there is limited guidance for those developing or researching health-related VR environments. Often, usability evaluation approaches used for other health information technology applications are difficult to implement within VR contexts. Thus, VR applications used in health contexts may not always undergo a thorough usability assessment. In the meantime, however, methods developed outside of the VR context will continue to be used until the scientific approaches for assessing VR usability further develop and until methods from the VR gaming industry become commonplace in health technology–related research.

VR Usability Assessment Methods

Overview

The following section describes VR usability assessment methods that have been employed in past research. It is important to note that these methods may be hybridized and blended together to suit the goals of each unique evaluation and are not mutually exclusive. The approaches are described using a previously developed classification of usability methods in virtual environments developed by Bowman and colleagues in 2000 [27] and updated by Martens in 2016 [28]. These approaches include (1) cognitive or task walkthrough, (2) graphical evaluation, (3) post hoc questionnaire or interview, (4) physical performance evaluation, (5) user interface (UI) evaluation, and (6) heuristic evaluation.

Table 1 [14,21,29-36] summarizes key information related to each of the identified VR assessment approaches, including

some considerations for assessment requirements excluding basic needs, such as an appropriate space to conduct a VR assessment and the VR hardware and software itself.

It is recommended that some assessment methods should favor the involvement of specific user groups, such as external users (ie, a group of testers not involved in the development process). Some assessment method requirements also lend themselves to requiring representative users, meaning a sample of users who may reflect the appropriate end-user population. The following sections provide an explanation of each of the VR usability assessment approaches.

Table 1. Overview of virtual reality (VR) usability assessment approaches.

Approach	Assessment requirements[a]	VR aspects evaluated	Typical output of assessment	Results type
Cognitive or task walkthrough [14,29]	Representative external users, developed task or scenario, and recording and timing equipment	Environment navigation, object interaction, and user-system interaction	Task performance and user feedback	Discrete and descriptive
Graphical evaluation [30,31]	Multiple relevant graphical environments, *recording equipment, questionnaires,* and *interview guides*	Quality of graphics and image renderings	User feedback	Descriptive
Post hoc questionnaires and interviews [32]	External users, questionnaires or interview guides, and *recording equipment*	Nonspecific	User feedback	Descriptive
Physical performance evaluation [33]	External users, developed task or scenario, and *recording and timing equipment*	Physical immersion and VR performance	Task performance, system performance metrics, and user feedback	Discrete and descriptive
User interface evaluation [14,34,35]	Developed task or scenario, *recording and timing equipment, questionnaires,* and *interview guides*	Integration of VR environment and real-life tools and VR performance	User feedback and task performance	Discrete and descriptive
Heuristic evaluation [14,21,36]	Experienced users, developed task or scenario, questionnaires, interview guides, and *recording equipment*	Various	Refer to list in Heuristic Evaluation section	Descriptive

[a]Italicized text indicates optional requirements depending on specific assessment approaches being used (eg, recording equipment is only required if incorporating think-aloud methods).

Cognitive or Task Walkthrough

The cognitive or task walkthrough is a formative assessment method that assesses the user, or hypothetical user, based on the completion of task-based VR scenarios, response to system changes, and the user's exploration and navigation of the VR environment [14]. While other measures for task load performance exist, such as the NASA-TLX (Task Load Index) [37], this assessment is based on Norman's 1986 [38] model of interaction and assesses the user's mental and physical actions in VR environments founded on the premise that users learn to use a technology through a process of self-exploration rather than didactic training or lessons [39]. Originally designed to assess simple UIs, such as automated teller machines and kiosks, this assessment method is increasingly used to assess VR usability as well [40].

One way to perform such an assessment is by employing the following three cycles or steps. The first cycle assesses a user's actions when they are trying to achieve a goal [14]. An observer will document the overall path the user takes to complete a task or whether they behave in an *intended* way in the VR scenario. Challenges or issues in achieving the goal of each cycle are noted by the observer. Behaviors in this first cycle are largely dictated by the user having to make decisions and how the environment facilitates such decision pathways. For example, if the user's goal is to pick up an object, but the object is missing, then the environment should allow the user to locate the object. Locating the object itself leads to the second cycle

or step called "exploration and navigation in virtual environments" [14].

In the second cycle, the user explores and moves around the environment to identify a path toward an object of interest. The VR environment should allow for intuitive navigation, recognizing user movements and responsively adapting to changes in user location as the user explores to locate the object of interest [14]. The observer records any challenges or issues in achieving this goal.

In the third cycle, the user's behaviors in response to a system initiative are assessed [14]. The purpose of this cycle or step is to examine how the VR system supports user activity when the user manipulates an object. The user and system are required to reciprocally recognize and interpret the feedback or actions of one another and respond appropriately [14]. For instance, if the user decides to throw a vase, the system should interpret this action and produce an appropriate response, such as depicting the vase flying and being shattered when contacting another object such as a wall. Correspondingly, the system may also take the initiative and act, meaning it is the user's role to interpret and respond to this action [14]. For example, if a helium balloon (ie, the object) suddenly detaches from its base and starts floating away (ie, the system's action or initiative), the user in this case may then choose to intervene and attempt to catch the balloon or allow the balloon to float away.

In summary, a cognitive or task walkthrough is a task-based assessment that assesses a user's actions when they are trying

to achieve a goal (see Table 1) and incorporates further assessments of user navigation (ie, cycle two) and system response (ie, cycle three). For each of these cycles, users should be allowed to freely walk through the task or interaction without interruption by the observers. Since this approach is largely driven by scripts and dialogues within the VR environment, usability issues and the system's ability to support user interaction is primarily assessed through descriptive, qualitative feedback (eg, user comments, think-aloud method, and observer observations) [14,28].

Graphical Evaluation

This assessment method focuses on the quality of graphics generated in the VR environment and how it influences the user's experience. This may include, but is not limited to, how different color combinations, shapes, textures, and renderings depicted will impact the user's interaction with the VR environment and system [30,41]. There are numerous ways to assess graphics, which can be attuned to hardware (eg, view, resolution, color contrast, update rate, etc); fidelity (eg, geometry and colors); camera placement, if applicable; the precision of the tracking system; stereoscopic image quality [42]; and beyond [43].

Many methods that vary in degrees of complexity exist for assessing graphics. In one common approach, users are exposed to different iterations of graphical environments to get a better understanding of its impact on user experience [30,41]. Depending on the purpose of the overall VR environment, the graphical object of interest may vary. For example, to examine a user's behavior in a large city, the graphical evaluation may be more focused on image depth, complexity, and breadth of the city and its 3D renderings. If the focus is narrower, such as assessing how a user reacts to seeing smoking paraphernalia, then focusing on meticulous, realistic details for an object such as a cigarette will be of greater importance. To assess a user's response to graphics in a VR environment, users may be asked to think aloud or be given a set of questionnaires to collect user feedback about the graphical output in the VR system (see Post Hoc Questionnaires and Interviews section) [30,41].

Post Hoc Questionnaires and Interviews

Post hoc questionnaires and interviews are often used to identify a user's general overall experience in using VR. However, some of these questionnaires and interviews may be targeted toward specific usability concepts, such as graphics, the physical hardware, and motion sickness [37,44,45].

This assessment method is often performed following the conclusion of a user's interaction with the VR system [28]. Since VR remains a relatively new technology, responses may be highly influenced by the individual's comfort and experience with using VR. Thus, unless the usability evaluation is already tailored to a target or only includes a subset of users based on experience (eg, inexperienced VR users), demographic information about users' opinions, views, and experiences with VR should also be collected to help better interpret user feedback [28]. Due to its versatility, this assessment can be viewed as a complement to many of the assessments covered in this article rather than a stand-alone method. Its overall

purpose is to serve as a simple, straightforward way of collecting targeted feedback. In order to collect specific feedback pertaining to the specific evaluation tied to a post hoc questionnaire or interview, special care must be given to the semantics and framing of questions [28].

Often, post hoc questionnaires are also used during the VR prototyping stage by engaging end users as a form of iterative quality improvement, but often in conjunction with another evaluation method such as a cognitive or task walkthrough [28].

Physical Performance Evaluation

Physical performance in the context of VR is defined by the performance of the hardware and environment. The smoothness and quality of the virtual environment are evaluated not unlike how a website can be evaluated on its loading time. Performance metrics with this assessment method include lag time (ie, the time delay between the user's intended action and the system's response within VR) and synchronization (ie, whether the system accurately reflects the user's intended actions). VR should be as convincingly realistic as possible to users, and the physical performance of a VR system is the key determinant of mental and physical immersion [33]. Immersion is defined as a state of being fully absorbed and/or deeply engaged within a simulated environment and is a key factor in determining the quality of VR [21]. This assessment method can facilitate user-centered design and can also yield information on the physical space required for users to fully explore the VR environment [33].

To gauge the VR system's physical performance, data can be obtained through a combination of approaches, such as questionnaires, task performance scores, or by leveraging back-end data to examine factors such as retrieval and load times. Simple but physically demanding VR precision tasks are highly informative for this type of assessment. For example, a task involving manipulating small objects with virtual chopsticks will quickly reveal any performance issues related to the precision of translated movements. Such a task can be timed and scored, and the user can be asked to describe their satisfaction and feelings to identify physical performance issues [33]. As another example, tasks involving actions that require users to reach out around their *body* to interact with nearby objects can be used to highlight unaddressed issues with distance compression, a frequent phenomenon within VR environments where objects are perceived by the user to be closer than their actual position [46]. Following a given task, a user may achieve high task performance scores but still report heavy cognitive overload (ie, mental exhaustion) while using the system, for example, finding that performing the task in VR was significantly more difficult than performing the same task with real objects or tools. Such a situation would signal that some probing questions (eg, Was there a specific action of the task that was particularly difficult to perform?) or further back-end evaluations may be required to identify possible underlying physical performance–related issues [33].

User Interface Evaluation

The purpose of a UI evaluation is to help determine the usability of a VR system's front-end UI [14,35,47]. This approach can

also help identify a UI design or solution that appropriately balances factors such as intuition and immersion against usability [34]. An optimized UI solution should provide the user with the best combination between immersion and usability, such that users feel immersed but unencumbered in accomplishing their tasks relative to outside a VR environment [14,34]. A feeling of immersion is especially pertinent when considering VR applications that notably outperform real-world counterparts, such as a simulated environment used to manage phobias or pain [8]. In these unique situations where the UI itself is deeply interrelated with the intervention (ie, phobia exposure tool), a comprehensive UI design evaluation may only be feasibly accomplished by a wider-scale clinical trial measuring treatment outcomes. Returning to more general VR applications, a UI evaluation allows for the identification of the type of UI solution that will provide the best immersion-to-efficiency ratio between a VR environment and real-life tools [31]. In a proof-of-concept case study by Kasurinen [34], users were instructed to complete one of five training scenarios with three varying levels of VR and real-life tools [34]:

1. No VR: participants move throughout an environment with keyboard and mouse controls; other activities are completed with real-life tools in a simulated workspace setting.

2. Semi-VR: participants move within a virtual environment with a VR headset; other activities are completed with real-life tools in a simulated workspace setting. The real-life workstation also displays the current state of the VR.

3. Full VR: participants move and complete their activities fully within a VR environment. Real-life tools are replaced with virtual equivalents (eg, virtual keyboard) and other real-life displays (eg, workstation screen) are virtually broadcasted to the VR headset.

For each iteration, data on user preferences can be collected alongside discrete data, such as task completion times and the number of errors [34]. Questions related to UI elements should also be asked throughout each iteration, as follows [14]:

1. Can the user form or remember the task goal?
2. Are the appropriate objects or parts of the environment viable?
3. Can the necessary objects be located?
4. Can the user execute movement and navigation actions?
5. Can the user recognize objects in the environment?

Each of these questions can help to reveal a specific area with potential for improvement within the UI. This method can aid in assessing both the appropriate amount of real-life integration and the quality of said integration so the VR intervention can best accomplish its intended purpose. If the integration between virtual and real-life tools is insufficient, it has been shown that this friction will cause users to prefer the *No VR* option, which may also be partially related to physical performance (see Physical Performance Evaluation section) [34].

Heuristic Evaluation

A heuristic evaluation is a UI approach that involves several topic experts or an expert evaluator, rather than soliciting direct user feedback. A VR usability expert will typically evaluate a UI's design against an accepted set of usability principles or standards already published in the literature [48]. While there are several sets of accepted standards or heuristics, for traditional UIs little research exists on defining heuristics for VR environments. Nielsen's [21] heuristics set is the most commonly referenced and utilized set of heuristics for UI design. Sutcliffe and Gault [48] further defined a set of 12 heuristic guidelines based on Nielsen's set, as shown in Textbox 1.

Textbox 1. A set of 12 heuristic guidelines.

- Natural engagement
- Compatibility with the user's task and domain
- Natural expression of action
- Close coordination of action and representation
- Realistic feedback
- Faithful viewpoints
- Navigation and orientation support
- Clear entry and exit points
- Consistent departures
- Support for learning
- Clear turn taking
- Sense of presence

Expert results are then aggregated and used to identify priority areas of action [28]. Heuristic assessments also require a set of tasks for the experts to experience. The nature of these tasks and the VR environments themselves should also be subjectively considered when carrying out a heuristic assessment, given the lack of standardization between various types of VR equipment

and software [28]. While not all heuristics may apply to a given VR application, such an evaluation has great potential to glean a rich, overall picture of the state of the application. For example, if the VR application is intended to be designed in a way that the user is automatically placed in an "inescapable" environment, then there is no relevance in assessing clear entry

and exit points (ie, the eighth heuristic guideline, *clear entry and exit points*) [28]. Since heuristics are broad rules of thumb rather than specific guidelines, they should not be treated as binary checkboxes, but rather as individual continuums that can each be an area for improvement, although binary elements may exist within. To illustrate, perhaps the heuristic guideline of *realistic feedback* is of particular interest, which outlines that the VR application should help users effectively recognize and recover from errors [21]. The presence or absence of a feature such as, for example, tangible error messages would constitute a binary checkbox, but the palatability and effectiveness of said error messages would be of higher importance. Is the problem or error precisely and concisely indicated? Is a potential solution

suggested? Is the language user friendly and free of codes or abbreviations, such as "A 50 (0x32) error has occurred"? Ultimately, considering and tracking multiple granular elements within each heuristic will aid greatly in obtaining actionable results to direct improvement.

Resources Where Additional Information on the Topic Can Be Found

Table 2 [14,21,23,29-34,36,37,39,40,42-45,49-54] provides a list of references specific to each of the approaches where readers can access additional information.

Table 2. Additional resources.

Assessment approach	References and resources
Cognitive or task walkthrough	[14,29,39,40,49,50]
Graphical evaluation	[30,31,42,43]
Post hoc questionnaires and interviews[a]	[23,32,37,44,45,51]
Physical performance evaluation	[33]
User interface evaluation	[14,34]
Heuristic evaluation	[14,21,36,52]
Other	[53,54]

[a]This represents a sample of many that can be employed, depending on what usability concept is to be measured.

Other Considerations

As previously noted, many of the approaches presented in this paper can be blended or hybridized together to suit the goals or needs of a given VR application evaluation. They are certainly not mutually exclusive. Some of the methods already incorporate a level of hybridization, most often with the inclusion of a post hoc questionnaire or interview. Given the lack of standardization across approaches, this warrants future research regarding the development of a comprehensive framework incorporating multiple methods of VR evaluation to provide, at a minimum, a strategic work plan for those looking to perform a baseline evaluation of any new VR application. This should include the incorporation of more up-to-date methods already used in the gaming industry. Those who employ usability methods for VR that have been developed for other kinds of health information technologies should be encouraged to share their experiences with the broader scientific community, placing an emphasis on the practical experiences of doing so. The current literature base lacks practical examples of how to best use these approaches, which could be of great use to those employing them.

When developing VR interventions and applications, particularly in the context of health, the comfort of the end user is paramount. Alongside the numerous benefits of VR technology, VR still carries the risk of imposing symptoms similar to motion sickness during use as a result of visual distortions and asynchronies, among other effects [45]. While these issues are peripherally related to performance issues and may be identified in user feedback, these data are inherently subjective and the effects are, thus, not easily quantifiable enough to measure

improvements. Thus, the authors recommend that any VR assessment also explicitly consider the possible effect of motion sickness on its users by incorporating tools such as the Simulator Sickness Questionnaire (SSQ), originally developed to help measure motion sickness for pilots in flight simulators [45]. The results from the SSQ or another similar questionnaire may identify specific considerations for certain populations, age groups, diagnoses, and beyond. Additionally, the repeated occurrence of specific symptoms or combinations of such from the SSQ (eg, eyestrain, nausea, and vertigo) can provide additional direction in identifying the root issues within the VR software and hardware [45].

Conclusions

Health-related applications using VR are a rapidly advancing area of development. Like all emerging technologies in health care, there is a need to ensure the quality and safety of these novel tools [55]. For VR, validated usability and assessment approaches are an important step before its deployment in real-world clinical settings. The assessment methods described here give developers and researchers a high-level overview of important elements to consider regarding the usability of their VR implementations and to make iterative changes prior to clinical implementation. However, once these approaches are employed for VR, sharing practical experiences in doing so would be of tremendous value. This area of science is in its infancy and comprehensive knowledge translation would be critical to its growth.

Overall, this paper provides a description and discussion of six different contemporary VR usability assessment methods. As

XSL·FO

RenderX

an emerging area for research, the development of formative usability assessment methodologies for health-related VR applications is an important area for future development. Further, while the six approaches discussed in this paper have been discussed in isolation, further future hybridization of approaches to develop more robust and multidimensional interpretations of VR usability should be considered. For instance, like other usability evaluation approaches [21,56], a purposeful mixed methods approach may assist in generating more holistic and robust interpretations of a system's usability. We see value in the triangulation of data related to user feedback and other task performance metrics in health-related VR applications. Due to the nascent nature of the domain, a pluralistic approach to usability evaluation should be considered in an effort to develop broader and more nuanced understandings of the state of the art in VR.

Given that the VR industry is projected to grow to over US $9 billion in sales of VR devices alone by 2021 [57], it is no surprise the industry is marked with large financial investments, such as the acquisition of Oculus for US $2 billion, as many large technology companies continue to invest heavily in VR [58]. As a collective, health care organizations and professionals should emphasize ensuring the mitigation and prevention of potential growing pains that may arise if VR interventions are churned out without rigorous evaluation and proper regard for quality, allowing for VR to usher in a new field of innovative, technology-enabled health care. With this foundation, the potential benefits to providers and patients alike will only continue to grow with continuous improvements in technology and reductions in cost.

Acknowledgments

Authors' Contributions

This work was first conceived by GS. All authors contributed to the writing and editing of the manuscript through feedback and discussion, met ICMJE (International Committee of Medical Journal Editors) author requirements, and have approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Hu J, Luo E, Song E, Xu X, Tan H, Zhao Y, et al. Patients' attitudes towards online dental information and a web-based virtual reality program for clinical dentistry: A pilot investigation in China. Int J Med Inform 2009 Mar;78(3):208-215. [doi: 10.1016/j.ijmedinf.2008.07.003] [Medline: 18752987]

2. Tossavainen T, Juhola M, Pyykkö I, Aalto H, Toppila E. Development of virtual reality stimuli for force platform posturography. Int J Med Inform 2003 Jul;70(2-3):277-283. [doi: 10.1016/s1386-5056(03)00034-0]

3. Vogt S, Skjæret-Maroni N, Neuhaus D, Baumeister J. Virtual reality interventions for balance prevention and rehabilitation after musculoskeletal lower limb impairments in young up to middle-aged adults: A comprehensive review on used technology, balance outcome measures and observed effects. Int J Med Inform 2019 Jun;126:46-58. [doi: 10.1016/j.ijmedinf.2019.03.009] [Medline: 31029263]

4. Hone-Blanchet A, Wensing T, Fecteau S. The use of virtual reality in craving assessment and cue-exposure therapy in substance use disorders. Front Hum Neurosci 2014;8:844 [FREE Full text] [doi: 10.3389/fnhum.2014.00844] [Medline: 25368571]

5. Bordnick PS, Graap KM, Copp HL, Brooks J, Ferrer M. Virtual reality cue reactivity assessment in cigarette smokers. Cyberpsychol Behav 2005 Oct;8(5):487-492. [doi: 10.1089/cpb.2005.8.487] [Medline: 16232041]

6. Clus D, Larsen ME, Lemey C, Berrouiguet S. The use of virtual reality in patients with eating disorders: Systematic review. J Med Internet Res 2018 Apr 27;20(4):e157 [FREE Full text] [doi: 10.2196/jmir.7898] [Medline: 29703715]

7. Galvez J, Eisenhower M, England W, Wartman E, Simpao A, Rehman M, et al. An interactive virtual reality tour for adolescents receiving proton radiation therapy: Proof-of-concept study. JMIR Perioper Med 2019 Mar 05;2(1):e11259 [FREE Full text] [doi: 10.2196/11259]

8. ClinicalTrials.gov. Bethesda, MD: US National Library of Medicine URL: https://clinicaltrials.gov/ [accessed 2020-10-06]

9. Gershon J, Zimand E, Lemos R, Olasov Rothbaum B, Hodges L. Use of virtual reality as a distractor for painful procedures in a patient with pediatric cancer: A case study. Cyberpsychol Behav 2003 Dec;6(6):657-661. [doi: 10.1089/109493103322725450] [Medline: 14756933]

10. Freitas DM, Spadoni VS. Is virtual reality useful for pain management in patients who undergo medical procedures? Einstein (Sao Paulo) 2019;17(2):1-3 [FREE Full text] [doi: 10.31744/einstein_journal/2019md4837]

11. Eijlers R, Utens EMWJ, Staals LM, de Nijs PFA, Berghmans JM, Wijnen RMH, et al. Systematic review and meta-analysis of virtual reality in pediatrics. Anesth Analg 2019;129(5):1344-1353. [doi: 10.1213/ane.0000000000004165]

12. Tarrant J, Viczko J, Cope H. Virtual reality for anxiety reduction demonstrated by quantitative EEG: A pilot study. Front Psychol 2018;9:1280 [FREE Full text] [doi: 10.3389/fpsyg.2018.01280] [Medline: 30087642]

13. Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: A randomized clinical trial. JAMA Psychiatry 2019 Jul 01;76(7):682-690 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.0219] [Medline: 30892564]

14. Sutcliffe AG, Kaur KD. Evaluating the usability of virtual reality user interfaces. Behav Inf Technol 2000 Jan;19(6):415-426. [doi: 10.1080/014492900750052679]

15. Nichols S, Patel H. Health and safety implications of virtual reality: A review of empirical evidence. Appl Ergon 2002 May;33(3):251-271. [doi: 10.1016/s0003-6870(02)00020-0] [Medline: 12164509]

16. Huygelier H, Schraepen B, van Ee R, Vanden Abeele V, Gillebert CR. Acceptance of immersive head-mounted virtual reality in older adults. Sci Rep 2019 Mar 14;9(1):4519 [FREE Full text] [doi: 10.1038/s41598-019-41200-6] [Medline: 30872760]

17. Han P, Chen Y, Lee K, Wang H, Hsieh C, Hsiao J, et al. Haptic around: Multiple tactile sensations for immersive environment and interaction in virtual reality. In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18). 2018 Presented at: 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18); November 28-December 1, 2018; Tokyo, Japan p. 1-10. [doi: 10.1145/3281505.3281507]

18. Riva G, Wiederhold BK, Mantovani F. Neuroscience of virtual reality: From virtual exposure to embodied medicine. Cyberpsychol Behav Soc Netw 2019 Jan;22(1):82-96 [FREE Full text] [doi: 10.1089/cyber.2017.29099.gri] [Medline: 30183347]

19. Walji MF, Kalenderian E, Piotrowski M, Tran D, Kookal KK, Tokede O, et al. Are three methods better than one? A comparative assessment of usability evaluation methods in an EHR. Int J Med Inform 2014 May;83(5):361-367 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.010] [Medline: 24583045]

20. Ellsworth MA, Dziadzko M, O'Horo JC, Farrell AM, Zhang J, Herasevich V. An appraisal of published usability evaluations of electronic health records via systematic review. J Am Med Inform Assoc 2017 Jan;24(1):218-226. [doi: 10.1093/jamia/ocw046] [Medline: 27107451]

21. Nielsen J. Enhancing the explanatory power of usability heuristics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94). 1994 Presented at: SIGCHI Conference on Human Factors in Computing Systems (CHI '94); April 24-28, 1994; Boston, MA p. 152-158.

22. Dey A, Billinghurst M, Lindeman RW, Swan JE. A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. Front Robot AI 2018 Apr 17;5:1-28 [FREE Full text] [doi: 10.3389/frobt.2018.00037]

23. Brooke J. SUS: A 'quick and dirty' usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry. London, UK: Taylor & Francis; 1996:189-194.

24. Saleem JJ, Haggstrom DA, Militello LG, Flanagan M, Kiess CL, Arbuckle N, et al. Redesign of a computerized clinical reminder for colorectal cancer screening: A human-computer interaction evaluation. BMC Med Inform Decis Mak 2011 Nov 29;11:74 [FREE Full text] [doi: 10.1186/1472-6947-11-74] [Medline: 22126324]

25. Johnson CM, Johnston D, Crowley PK, Culbertson H, Rippen HE, Damico DJ, et al. EHR Usability Toolkit: A Background Report on Usability and Electronic Health Records. Rockville, MD: Agency for Healthcare Research and Quality, US Department of Health and Human Services; 2011 Aug. URL: https://digital.ahrq.gov/sites/default/files/docs/citation/EHR_Usability_Toolkit_Background_Report.pdf [accessed 2020-10-06]

26. Karahoca A, Bayraktar E, Tatoglu E, Karahoca D. Information system design for a hospital emergency department: A usability analysis of software prototypes. J Biomed Inform 2010 Apr;43(2):224-232 [FREE Full text] [doi: 10.1016/j.jbi.2009.09.002] [Medline: 19755173]

27. Bowman DA, Gabbard JL, Hix D. A survey of usability evaluation in virtual environments: Classification and comparison of methods. Presence 2002 Aug;11(4):404-424 [FREE Full text] [doi: 10.1162/105474602760204309]

28. Martens D. Virtually Usable: A Review of Virtual Reality Usability Evaluation Methods [master's thesis]. New York, NY: Parsons School of Design; 2016 May 14. URL: https://danamartensmfadt.files.wordpress.com/2016/08/virtuallyusable.pdf [accessed 2020-10-06]

29. Costalli F, Marucci L, Mori G, Paterno F. Design criteria for usable web-accessible virtual environments. In: Proceedings of the International Cultural Heritage Informatics Meeting (ICHIM 2001). 2001 Presented at: International Cultural Heritage Informatics Meeting (ICHIM 2001); September 3-7, 2001; Milan, Italy p. 413-426.

30. McMahan RP, Bowman DA, Zielinski DJ, Brady RB. Evaluating display fidelity and interaction fidelity in a virtual reality game. IEEE Trans Vis Comput Graph 2012 Apr;18(4):626-633. [doi: 10.1109/tvcg.2012.43]

31. Vaden EA, Ehrlich JA, Kolasinski EM. Usability evaluation of low-end virtual reality systems. In: Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting. 1996 Presented at: Human Factors and Ergonomics Society 40th Annual Meeting; September 2-6, 1996; Philadelphia, PA. [doi: 10.1177/154193129604002481]

32.   Castilla D, Garcia-Palacios A, Bretón-López J, Miralles I, Baños RM, Etchemendy E, et al. Process of design and usability evaluation of a telepsychology web and virtual reality system for the elderly: Butler. Int J Hum Comput Stud 2013 Mar;71(3):350-362. [doi: 10.1016/j.ijhcs.2012.10.017]

33.   Rezazadeh IM, Firoozabadi M, Wang X. Evaluating the usability of virtual environment by employing affective measures. In: Wang X, editor. Mixed Reality and Human-Robot Interaction. Intelligent Systems, Control and Automation: Science and Engineering, vol 1010. Dordrecht, the Netherlands: Springer; 2011:95-109.

34.   Kasurinen J. Usability issues of virtual reality learning simulator in healthcare and cybersecurity. Procedia Comput Sci 2017;119:341-349. [doi: 10.1016/j.procs.2017.11.193]

35.   Chin J, Diehl VA, Norman KL. Development of an instrument measuring user satisfaction of the human-computer interface. In: Proceedings of the Human Factors in Computing Systems Conference (ACM CHI '88). 1988 Presented at: Human Factors in Computing Systems Conference (ACM CHI '88); May 15-19, 1988; Washington, DC p. 213-218. [doi: 10.1145/57167.57203]

36.   Sutcliffe AG, Deol Kaur K. A usability evaluation method for virtual reality user interfaces. CiteSeerX. 2006. URL: http:/ /citeseerx.ist.psu.edu/viewdoc/download;jsessionid=B5295890A3F6C06F50E9FC12DF1EB4F3?doi=10.1.1.115. 945&rep=rep1&type=pdf [accessed 2020-10-06]

37.   Hart S, Staveland L. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Adv Psychol 1988;52:139-183. [doi: 10.1016/S0166-4115(08)62386-9]

38.   Norman DA. Cognitive engineering—Cognitive science. In: Carroll JM, editor. Interfacing Thought: Cognitive Aspects of Human-Computer Interaction. Cambridge, MA: MIT Press; 1987:325-336.

39.   Polson P, Lewis C, Rieman J, Wharton C. Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. Int J Man Mach Stud 1992 May;36(5):741-773. [doi: 10.1016/0020-7373(92)90039-N]

40.   Wharton C, Bradford J, Jeffries R, Franzke M. Applying cognitive walkthroughs to more complex user interfaces experiences, issues, and recommendations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92). 1992 Presented at: SIGCHI Conference on Human Factors in Computing Systems (CHI '92); May 3-7, 1992; Monterey, CA p. 381-388. [doi: 10.1145/142750.142982]

41.   Van Orden KF, DiVita J. Effects of structure and color on symbol visibility. In: Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting. 1996 Presented at: Human Factors and Ergonomics Society 40th Annual Meeting; September 2-6, 1996; Philadelphia, PA. [doi: 10.1177/154193129604002482]

42.   Moorthy A, Su C, Mittal A, Bovik A. Subjective evaluation of stereoscopic image quality. Signal Process Image Commun 2013 Sep;28(8):870-883. [doi: 10.1016/j.image.2012.08.004]

43.   Menzies RJ, Rogers SJ, Phillips AM, Chiarovano E, de Waele C, Verstraten FAJ, et al. An objective measure for the visual fidelity of virtual reality and the risks of falls in a virtual environment. Virtual Real 2016 Jun 6;20(3):173-181. [doi: 10.1007/s10055-016-0288-6]

44.   Golding J. Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. Brain Res Bull 1998 Nov 15;47(5):507-516. [doi: 10.1016/s0361-9230(98)00091-4] [Medline: 10052582]

45.   Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. Int J Aviat Psychol 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]

46.   Finnegan D, O'Neill E, Proulx M. An approach to reducing distance compression in audiovisual virtual environments. In: Proceedings of the IEEE 3rd VR Workshop on Sonic Interactions for Virtual Environments (SIVE). 2017 Presented at: IEEE 3rd VR Workshop on Sonic Interactions for Virtual Environments (SIVE); March 19, 2017; Los Angeles, CA p. 1-6. [doi: 10.1109/SIVE.2017.7901607]

47.   Charfi S, Kolski C. RITA: A useR Interface evaluaTion frAmework. J Univers Comput Sci 2015;21(4):526-560 [FREE Full text] [doi: 10.3217/jucs-021-04-0526]

48.   Sutcliffe A, Gault B. Heuristic evaluation of virtual reality applications. Interact Comput 2004 Aug;16(4):831-849. [doi: 10.1016/j.intcom.2004.05.001]

49.   Wilson C. Cognitive walkthrough. In: User Interface Inspection Methods: A User-Centered Design Method. Waltham, MA: Morgan Kaufmann; Nov 2013:65-80.

50.   Mahatody T, Sagar M, Kolski C. State of the art on the cognitive walkthrough method, its variants and evolutions. Int J Hum Comput Interact 2010 Jul 30;26(8):741-785. [doi: 10.1080/10447311003781409]

51.   Golding JF. Predicting individual differences in motion sickness susceptibility by questionnaire. Pers Individ Dif 2006 Jul;41(2):237-248. [doi: 10.1016/j.paid.2006.01.012]

52.   Murtza R, Monroe S, Youmans R. Heuristic evaluation for virtual reality systems. In: Proceedings of the Human Factors and Ergonomics Society International Annual Meeting. 2017 Presented at: Human Factors and Ergonomics Society International Annual Meeting; October 9-13, 2017; Austin, TX p. 2067-2071. [doi: 10.1177/1541931213602000]

53.   Dias P, Pimentel A, Ferreira C, Madeira J. Usability in virtual and augmented environments: A qualitative and quantitative study. In: Proceedings of the International Society for Optics and Photonics (SPIE): Stereoscopic Displays and Virtual Reality Systems XIV. 2007 Mar Presented at: International Society for Optics and Photonics (SPIE): Stereoscopic Displays and Virtual Reality Systems XIV; January 29-31, 2007; San Jose, CA. [doi: 10.1117/12.703878]

54.   Livatino S, Koffel C. Handbook for evaluation studies in virtual reality. In: Proceedings of the IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems. 2007 Presented at: IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems; June 25-27, 2007; Ostuni, Italy p. 1-6. [doi: 10.1109/VECIMS.2007.4373918]

55.   Mytton OT, Velazquez A, Banken R, Mathew JL, Ikonen TS, Taylor K, et al. Introducing new technology safely. Qual Saf Health Care 2010 Aug;19 Suppl 2:i9-i14. [doi: 10.1136/qshc.2009.038554] [Medline: 20693217]

56.   Xiao Y, Montgomery DC, Philpot LM, Barnes SA, Compton J, Kennerly D. Development of a tool to measure user experience following electronic health record implementation. J Nurs Adm 2014;44(7/8):423-428. [doi: 10.1097/nna.0000000000000093]

57.   Flavián C, Ibáñez-Sánchez S, Orús C. The impact of virtual, augmented and mixed reality technologies on the customer experience. J Bus Res 2019 Jul;100:547-560. [doi: 10.1016/j.jbusres.2018.10.050]

58.   Cipresso P, Giglioli IAC, Raya MA, Riva G. The past, present, and future of virtual and augmented reality research: A network and cluster analysis of the literature. Front Psychol 2018;9:2086 [FREE Full text] [doi: 10.3389/fpsyg.2018.02086] [Medline: 30459681]

Abbreviations

ICMJE: International Committee of Medical Journal Editors
SSQ: Simulator Sickness Questionnaire
TLX: Task Load Index
UI: user interface
VR: virtual reality

XSL•FO
RenderX

Review

Serious Gaming Technology in Upper Extremity Rehabilitation: Scoping Review

Elisavet Koutsiana[1], MSc; Ioannis Ladakis[1], MSc; Dimitris Fotopoulos[1], MSc; Achilleas Chytas[1], MSc; Vassilis Kilintzis[1], PhD; Ioanna Chouvarda[1], PhD

Lab of Computing, Medical Informatics, and Biomedical-Imaging Technologies, School of Medicine, Aristotle University, Thessaloniki, Greece

Corresponding Author:
Elisavet Koutsiana, MSc
Lab of Computing, Medical Informatics, and Biomedical-Imaging Technologies
School of Medicine
Aristotle University
Building 16d, 2nd floor
Thessaloniki, 54124
Greece
Phone: 30 2310999247
Email: elisavetkoutsiana@gmail.com

Abstract

Background: Serious gaming has increasingly gained attention as a potential new component in clinical practice. Specifically, its use in the rehabilitation of motor dysfunctions has been intensively researched during the past three decades.

Objective: The aim of this scoping review was to evaluate the current role of serious games in upper extremity rehabilitation, and to identify common methods and practice as well as technology patterns. This objective was approached via the exploration of published research efforts over time.

Methods: The literature search, using the PubMed and Scopus databases, included articles published from 1999 to 2019. The eligibility criteria were (i) any form of game-based arm rehabilitation; (ii) published in a peer-reviewed journal or conference; (iii) introduce a game in an electronic format; (iv) published in English; and (v) not a review, meta-analysis, or conference abstract. The search strategy identified 169 relevant articles.

Results: The results indicated an increasing research trend in the domain of serious gaming deployment in upper extremity rehabilitation. Furthermore, differences regarding the number of publications and the game approach were noted between studies that used commercial devices in their rehabilitation systems and those that proposed a custom-made robotic arm, glove, or other devices for the connection and interaction with the game platform. A particularly relevant observation concerns the evaluation of the introduced systems. Although one-third of the studies evaluated their implementations with patients, in most cases, there is the need for a larger number of participants and better testing of the rehabilitation scheme efficiency over time. Most of the studies that included some form of assessment for the introduced rehabilitation game mentioned user experience as one of the factors considered for evaluation of the system. Besides user experience assessment, the most common evaluation method involving patients was the use of standard medical tests. Finally, a few studies attempted to extract game features to introduce quantitative measurements for the evaluation of patient improvement.

Conclusions: This paper presents an overview of a significant research topic and highlights the current state of the field. Despite extensive attempts for the development of gamified rehabilitation systems, there is no definite answer as to whether a serious game is a favorable means for upper extremity functionality improvement; however, this certainly constitutes a supplementary means for motivation. The development of a unified performance quantification framework and more extensive experiments could generate richer evidence and contribute toward this direction.

XSL·FO
RenderX

Introduction

Serious Gaming in Upper Limb Motor Rehabilitation

Motor rehabilitation in various parts of the body such as the upper or lower limbs aims to help patients restore dysfunctions that affect their mobility. In this scoping review, we focus on motor disabilities related to the upper extremities. The motivation behind this review was first introduced within one of our group's research projects related to upper limb rehabilitation, termed "Modern Interface Platform for Motor Control and Learning on People With Motor Disorders" [1-3]. Our purpose was to search the literature regarding upper limb rehabilitation using serious games to provide guidance for proceeding with creation of the project's platform. With the term "serious games," we refer to video games created with a purpose other than entertainment, such as education, health care, politics, and engineering. The aim of this study was to review all of the upper limb rehabilitation techniques related to serious games regardless of the cause of motor dysfunctions.

Therapists have developed several clinical methods to indicate motor ability, such as range of motion (ROM) or range of force. In addition, specialized evaluation tests such as Fugl-Meyer Motor Function Assessment (FMA), Action Research Arm Test (ARAT), and Melbourne Assessment of Unilateral Upper Limb Function (MAUULF) aim to estimate the improvement of a patient's motion condition. The usual rehabilitation scheme consists of repeated motion exercises for a specific body part, with the aim of restoring ability as close to the normal condition as possible.

The idea to introduce gamification to the therapeutic protocol of upper limb rehabilitation was born as a means to motivate patients during the rehabilitation schemes but also represents a new method for monitoring the upper limb motion for further analysis. The first attempts of the introduction of gamification in upper limb rehabilitation appeared in 1999 by a team at Rutgers University [4], making use of a custom prototype robotic arm aiming to map the motion of the palm and wrist with force resistance. This concept was extended with development of a computer-based game that guides the patient to make various movements with the palm and fingers. The same system went through various modifications [5-7], and the latest version of the system was published a few years later [8-11], including significant alterations and improvements regarding the digital environment and the therapeutic approach. Among these early attempts, a study published in 2000 [12] presented a system that uses a robotic device in conjunction with the commercial game Arkanoid for wrist rehabilitation, and another study published in 2002 [13] described an equivalent approach using a resistive joystick.

These rapid technological developments led to more elaborate devices regarding motion capture, challenging researchers in this field to investigate this type of rehabilitation.

Significance of This Scoping Review

Over the last few decades, there has been an increasing amount of studies regarding the enhancement of rehabilitation with the introduction of new technologies. A systematic review on the implementation of serious games and wearable technology in rehabilitation practices for patients recovering from traumatic bone and soft tissue injuries was published by Meijer et al [14]. Another review attempted to depict the implementations of brain-computer interfaces in the rehabilitation of motor dysfunctions following stroke [15]. Nonetheless, these overviews do not include games specifically developed for rehabilitation or "wearable-controlled" games. Therefore, the primary aim of this scoping review was to summarize the field of upper extremity rehabilitation combined with serious games, providing a map of the research approaches used to date. The main research goals were to: (1) explore the technologies used for upper limb rehabilitation; (2) discover distinct methods, common characteristics, and objectives of these efforts; (3) identify challenges and limitations from these previous efforts; and (4) examine the types of analysis methods used to quantify the treatment outcome.

This effort will contribute to the detection of gaps or limitations in this area, and may lead to new research paths and ideas.

The rest of the paper is organized as follows. The Methods section depicts the procedure that was followed regarding the literature search, data management, and eligibility criteria of this review. The Results section presents the statistical results, including figures, after reviewing the included studies. Finally, the Discussion section comments on the results and delineates possible limitations of this study, along with highlighting the importance of this review for further development of this research area.

Methods

Design

In this scoping review, we followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analysis extension for Scoping Reviews) [16] guidelines for the literature search, study selection, and extracted information. We further referred to studies on scoping review methodology, including Arksey and O'Malley [17] and Peters et al [18].

Literature Search

This review included articles published from 1999 to June 2019. The PubMed and Scopus databases were used for the literature search. The keywords utilized in the literature search were: "rehabilitation," "hand," "upper limb," "upper extremity," "upper arm," "game," "serious gaming," and "serious game," which were investigated in titles and abstracts of articles published in the English language. The following search query was used: rehabilitation AND (hand OR upper limb OR upper-limb OR upper extremity OR upper-extremity OR upper arm OR upper-arm) AND (game OR serious gaming OR serious game). Subsequently, duplicated articles were removed, and the remaining studies were screened for eligibility.

Data Management

Two individual researchers (EK and IL) conducted the literature search and the removal of duplicates, and one author (IL) screened the titles and abstracts for eligibility under advisement by IC. The remaining studies were reviewed by EK, IL, and

DF, guided by a set of inclusion and exclusion criteria to extract information from the selected articles. The extracted information followed a structure defined by IC, EK, and IL, as follows: (i) year of publication, (ii) purpose of the study, (iii) part of the upper limb for rehabilitation, (iv) sensors used, (v) disease that led to the patients' condition, (vi) game type, (vii) game scenario, (viii) game target, (ix) clinic- or home-based application, (x) supervised or unsupervised, (xi) software used for creation of the rehabilitation game, (xii) hardware development, (xiii) system limitation, (xiv) use of a pilot study or not, (xv) number of patients in the pilot study, (xvi) evaluation methodology, and (xvii) features extracted from the game.

The literature search was conducted in July 2019 with the requirements described above, and a total of 682 studies were identified, including 151 from the PubMed database and 531 from the Scopus database. After removal of duplicates, 557 studies were screened with the criteria set, resulting in a total of 244 articles. In addition, 75 studies were excluded due to meeting one or more exclusion criteria, and 169 studies were finally included in the scoping review about upper limb rehabilitation based on serious gaming technology. The most common reasons for a study to be excluded were the absence of a serious game from the rehabilitation procedure and the development of a system that did not focus on upper extremity rehabilitation. Figure 1 shows the flow diagram of the exclusion stages for this review.

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analysis extension for Scoping Reviews) [16] flow diagram of the literature search and final included studies.
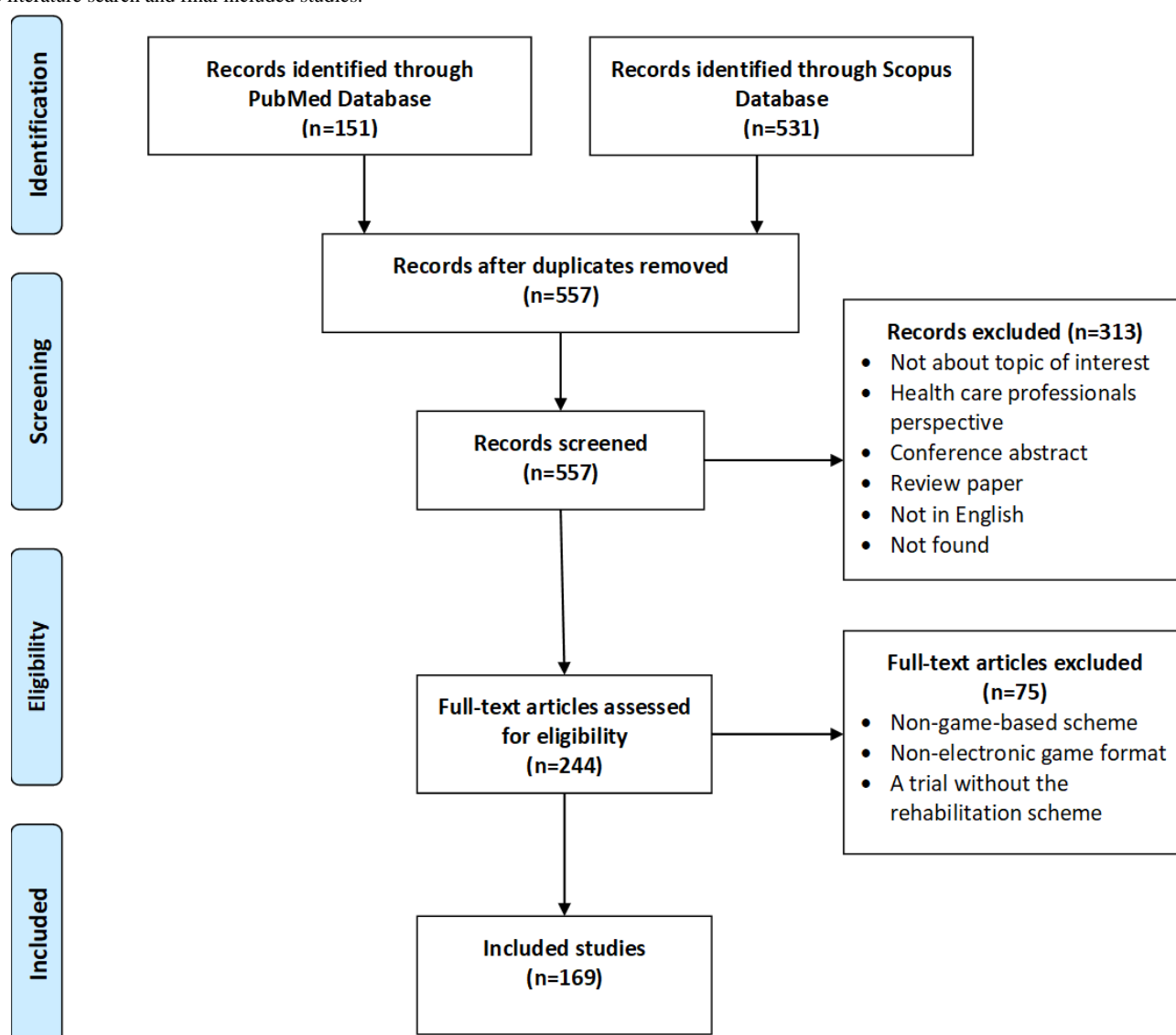
Inclusion and Exclusion Criteria

Inclusion criteria for eligibility of the selected articles were: (i) any form of game-based arm rehabilitation (interactive computer-based game, mobile/table app, or platform game software) and (ii) published in a peer-reviewed or conference journal. Exclusion criteria for this review were: (i) a nonserious game–based scheme of rehabilitation with a sensor (only using a sensor or robotic arm, without the accompanying serious game); (ii) a trial of a serious game rehabilitation scheme without any technical description of the game; (iii) medical article based on a health care professional's perspective for arm rehabilitation without any technical description of a game; (iv)

not published in English; or (v) a review, meta-analysis, or conference abstract.

Synthesis of Results

Based on the extracted information, we created 11 factors of categorization for the data. The extracted information is presented in the Data Management subsection above. The factors were determined based on the combination of the extracted information. All analysis factors were categorical, except for *device development*, which was a Boolean factor. In some cases, studies could belong to more than one category (eg, some studies mentioned analyses on both the wrist and fingers as targeted *upper extremity parts*, while others included both the score and time for the *game target*). Descriptive statistics were used for these factors to present an aggregated view of the studies and percentages.

The results extracted from the included studies are presented according to the following structure: (i) statistics depicted in charts, (ii) descriptive statistics that provide information

regarding the tendencies of research efforts, and (iii) conclusions extracted not only from the statistics but also from the general picture formed from the analysis of all included studies.

Results

Overview of Extracted Studies and Factors

Based on our literature search, the first study was published in 1999; however, only a few relevant papers were published in this field up to 2006. In 2007, researchers showed greater interest in upper extremity rehabilitation using new technologies based on serious games, and the number of publications has continued to rise up to the present day. Figure 2 summarizes the studies published on upper limb rehabilitation using serious games over the years. Notably, we only included studies published until June 2019, which means that the line graph in Figure 2 presents only half of the year for 2019. Table 1 summarizes the main factors that were used to draw conclusions and that were further analyzed.

Figure 2. Distribution of the publications over time. The "Commercial Sensors" category refers to studies using commercial sensors or any combination of commercial devices for the rehabilitation scenario, and the "Hardware Development" category refers to studies that created any type of robotic arm, glove, or other device for the connection with the game platform.
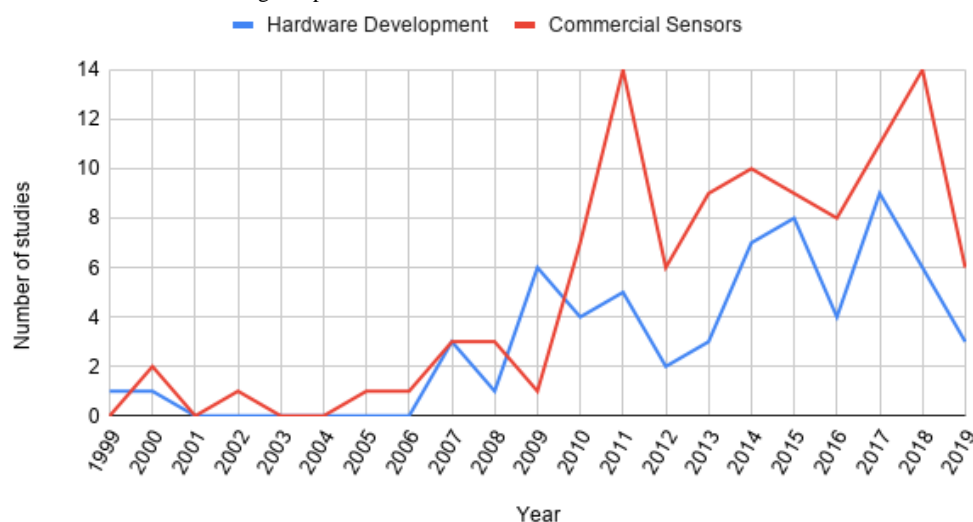
Table 1. Factors analyzed in the review.

Extracted information	Description
Medical condition	Underlying categories for the upper limb motion problems: stroke, general motion deficits, cerebral palsy, hemiparesis, neurological motor deficits, Parkinson disease, burn contractures, brain impairment, general cognitive deficits, shoulder injury, and wrist injury.
Upper extremity part	The upper extremity part targeted for rehabilitation: upper extremity/limb, fingers, palm, wrist, elbow, forearm, shoulder, and hand muscles.
Device development	The differentiation between hardware development and commercial sensors. The "commercial sensors" category includes studies that used commercial sensors/devices or any combination of commercial devices for the rehabilitation scenario, whereas the "hardware development" category includes studies that created any type of robotic arm, glove, or other device for the connection with the game platform
Game type	The type of serious game developed: virtual reality, augmented reality, video game, electronic board game, and mobile health apps.
Game target	The game scenario. Task completion (specific scenarios regarding the case, such as follow the line, daily activities, create shapes, collect a number of items), time (complete the level of the game in time intervals), score (increase the level score), force (studied the patient's force in controlling the sensor).
Sensors	The preferred device for the upper extremity exercise in conjunction with the game.
Hardware use limitations	Limitations regarding the hardware used in the included studies.
Supervision level	The rehabilitation scheme takes place at home or in a clinic while the patient is supervised by an expert or not.
System testing with users: pilots and trials	The testing part of the proposed scenario. Some studies conducted a trial or a pilot trial for testing the rehabilitation system, while others did not. In the latter studies, the trials were conducted with patients, control subjects, or both.
System evaluation	In cases in which the proposed system was tested, there were several means of evaluation: questionnaires, interviews, clinical tests, and scores before and after the rehabilitation scheme.
Extracted game features	The extracted characteristics, using the game, for further analysis: time, game performance, kinematic indicators, range of motion.

Medical Condition

To specify the *medical condition*, the categories were created based on the references used by the authors of the included papers regarding the medical condition that caused the motor dysfunction. For example, stroke and cerebral palsy are subcategories of hemiparesis or neurological motor deficits. However, some studies mentioned only hemiparesis or neurological motor deficits as a cause of motor dysfunction without any further explanation, while others specified that stroke was the cause of motor disability. Owing to this heterogeneity, we created the categories based on the references for the studied medical conditions. According to our literature search, stroke was the most common reason for upper extremity motor dysfunction and the need for rehabilitation using technology. More than half (56.8%, 96/169) of the studies introduced a system for upper limb rehabilitation after stroke, followed by general motion deficits (29.6%, 50/169 studies). Additional categories with lower frequency in the retrieved literature were cerebral palsy, hemiparesis, and neurological motor deficits. Furthermore, one study was related to Parkinson disease [19], one study was related to burn contractures [20],

two studies addressed patients with brain impairment [21,22], three focused on general cognitive deficits [23-25], and other conditions were identified as shoulder [26] or wrist [12,27] injuries.

Upper Extremity Part

Most of the included studies (48.5%, 83/169) referred to rehabilitation of the upper extremity/limb in general, whereas others (40.2%, 68/169) focused on a specific part such as the fingers, palm, wrist, elbow, forearm, and shoulder.

Device Development

The results of our search indicated that researchers in this field are showing more interest in commercial sensors that continue to evolve. Less interest is placed on the development of new devices designed for motion of a specific upper extremity part. This may be due to the more costly and time-consuming development of such specific devices. Studies on hardware development accounted for 37.3% (63/169) of the total studies, whereas there was double the number of studies related commercial sensors, representing 62.7% (106/169 studies) of the total. Despite the fewer attempts to address hardware

development, there seems to be continuous interest in this research area. However, it is evident that there are more fluctuations of publications over time in the case of hardware development due to the difficulty of the task (ie, the time and knowledge required to create a device), whereas commercial sensor–related studies showed a consistent increasing trend over time (Figure 2).

Between the two *device development* categories, the results regarding representation of the studied *medical condition* and *upper extremity part* factors did not vary in general. Stroke and upper extremity/limb accounted for more than half of all studies related to commercial sensors (56.6%, 60/106 and 48.1%, 51/106 studies, respectively) and hardware development (57%, 36/63 and 51%, 32/63, respectively). Nevertheless, these two categories have several differences concerning the *game type* and *game target* approaches, as described below.

Game Type

Game types were classified into different categories of virtual reality (VR), augmented reality (AR), video game, electronic board game, and mobile health (mHealth) apps. The included studies generally used the term "VR" to refer to games simulating the real world in 3D virtual environments generated by computer graphics (not necessarily using VR headsets), whereas the term "AR" is generally used to refer to two types of games: (1) games showing a real environment, but some objects are enhanced by computer-generated perceptual information; and (2) games representing a virtual world, including the real upper limb of the user (eg, using cameras).

Furthermore, "video games" refer to the creation of 2D games, whereas "electronic board games" refer to an interactive table or board. Finally, the "mHealth apps" category includes studies that describe the games as VR or AR, which are health apps using mobile or tablet games.

Based on these definitions, the majority of studies included in the review (74.6%, 126/169) approached the rehabilitation problem by developing VR games. This *game type* offers an alternative reality to the patient, transforming a repetitive exercise of a rehabilitation routine into an amusing and appealing game to spend their time.

The next most common *game type* was video games (15.4%, 26/169), along with some efforts to develop AR systems (10.1%, 17/169). With respect to AR, most of these studies used cameras and markers on the hand to recreate objects on the screen [28-39], although some recent studies used advanced technologies to create a 3D reality [23,40,41].

Comparing the two categories of *device development* regarding *game type*, the ratio of VR and video games was proportionally equal (Figure 3). Surprisingly, AR games exhibited essential differences in the two categories, with 7% (14/17) of the AR studies belonging to the category of commercial sensors and only 2% (3/17) belonging to the category of hardware development. In cases of AR, as mentioned above, most researchers used cameras and markers on the hand to capture the movement and incorporate it in the game (ie, combined commercial devices), which explains the higher percentage of studies in the commercial sensors category.

Figure 3. Comparison between the two *device development* categories regarding the *game type*. The ratio of virtual reality (VR), augmented reality (AR), video games, and electronic board games between the *commercial sensors* and *hardware development* categories is shown.
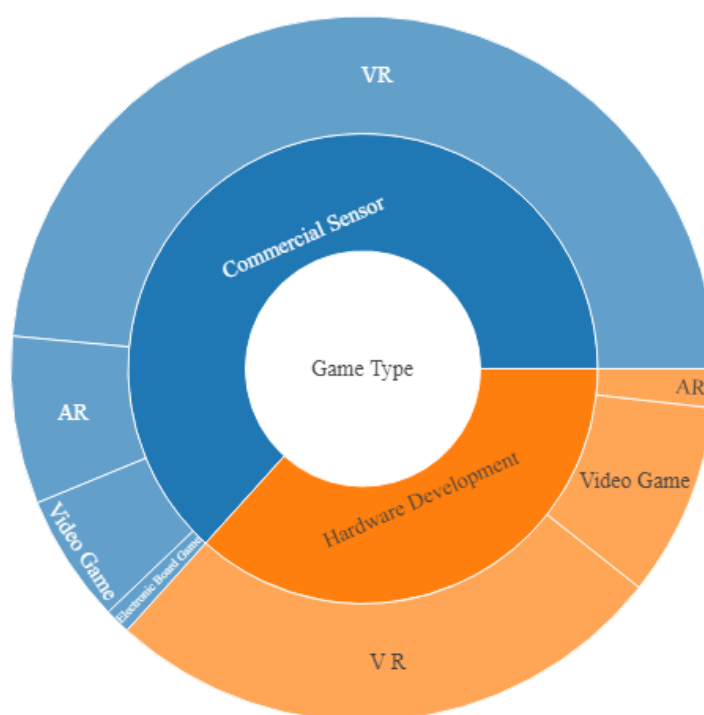
Over the years, technological development has led to increased incorporation of smartphones and tablets in our daily routines. Naturally, researchers have begun testing this new technology in many fields, including upper limb rehabilitation based on serious games by creating mHealth apps. Seven studies [35,40,42-46] created mobile games, and five studies [47-51] developed tablet games in an attempt to study a portable and easy-to-use-anywhere system for patients to perform their exercises. However, the touchscreen is the main component representing the evolution of board games. Two surveys [52,53] studied the use of electronic board games during supervised clinic sessions.

Game Target

The *game target* classification is summarized in Table 1. The highest percentage in this category was based on the target task completion (62.1%, 105/169 studies) and score (41.1%, 70/169 studies), with both designs focused on training to improve patients' ability. The approaches for the game scenario did not substantially differ between the two *device development* categories. Most games adopted scenarios similar to the corresponding therapy process and imitated movements from daily activities such as lifting a cup. Thus, the scenario varied depending on the part of the upper extremity that was targeted. Nevertheless, it is evident that in the case of attempts that belong to hardware development, the purpose is focused on more specific (fine) movements (ie, accuracy of the movement achieved in object placement), with noticeable interest in the force that the user exerts [13,47,54-61]. Moreover, another class of *game type* is the time (ie, completion time of the tasks). Overall, 16.6% (28/169) of the studies aimed at achieving time reduction of the specific task, thereby motivating the user to compete with themselves.

Sensors

Commercial Sensors

A variety of commercial or noncommercial devices have been proposed for rehabilitation of the upper extremities combined with serious gaming based on the researchers' ideas and accessible technologies at the time of publication. The most commonly used *commercial sensor* is the Kinect depth sensor, an accessory developed for the gaming platform Xbox. Kinect seems to be the most preferred sensor for capturing body parts and following their movement in space, which was used by 15.4% (26/169) of the studies included in the review. Some of the studies used only the Kinect sensor for their systems [26,62-73], whereas others combined it with biosignal capturing devices such as electromyogram (EMG) [24,41,74,75] or a sensing jacket [52] to gain better control of the user's movement for the final goal (ie, rehabilitation). In addition, some studies have used Kinect combined with gaming devices such as VR headsets [76] and a Wii balance board [77] or other devices such as goniometers [78-80], Tyromotion Timo plate [77], Xsen 3D sensor [81], body markers [82], and a customized haptic glove [83]. Furthermore, two studies focused on a different brand of depth sensor for their research, termed PrimeSense [84,85].

With respect to commercial gaming accessories, a few studies focused on individual sensors such as VR headsets [86-88], Wii remotes [89-95], or the P5 glove [96,97] in an attempt to incorporate the existing devices to rehabilitation practices. Another *commercial sensor* that has attracted researchers' interest is Leap Motion, a hand-tracking sensor, which is most commonly used alone [19,25,40,43,98-103]. One study also combined the Leap Motion sensor and a VR headset [104] in an attempt to create a VR environment for the user as a reinforcement of after-stroke rehabilitation methods. Another study [105] combined the Leap Motion sensor with a thermographic camera and a radiofrequency identification system for body part identification.

Additionally, some studies have attempted to create their own tracking system using *commercial sensors*. In some cases [57,58,78-80,106-110], sets of inertial measurement units (IMUs) were used as basic tracking sensors to measure the body's force, orientation, and angular rate. The sensors were placed on different parts of the upper limb or body to track the coordinates of the arm and, consequently, the arm movement. Some studies combined IMUs with Kinect to better determine the placement and movement of the body in space.

Overall, 17.8% (30/169) of the studies included webcams and cameras in their systems. Half of them [8,29-31,33,34,37,55,90,94,111-115] used only webcams in an attempt to create a home-based and easy-to-use patient system. The other half used either simple cameras [36] or cameras combined with a marker (ie, glove, card) to track the movement [21,39,116], gaming accessories such as PlayStation controllers [117] and Nintendo Wii remote [89], or a customized exoskeleton glove [118] and an eye tracker [38]. One study also included a motion-capture thermal camera [119] for motion detection in an attempt to avoid holding or wearing any controls or devices, which may be challenging and restrictive for the patient.

Besides interest in developing robotic devices identified in the hardware development category, some research teams have also focused on robotic devices that are already available on the market. Several studies used haptic devices such as Phantom Omni [22,120-122], Novint Falcon [77,122,123], Haptic Master [59,124-127], and Geomagic Touch [128], whereas others used robotic gloves such as CyberGrasp [124,126,129] and 5DT Data Glove [130,131] or robotic arms such as Barrett Wam [132] and Armeo Spring [20,108,133].

In addition, some studies attempted to either control or monitor patients' movements using medical devices such as EMG [30-32,34,41,44,87,112,134-142]. With EMG, it is possible to monitor how the muscles respond to nerve signals. In this way, physicians could observe the patient's upper extremity motion to prevent risky movements or to be sure that the patient is controlling the arm in the right direction based on his/her rehabilitation scheme. Moreover, some researchers have investigated the use of standard medical devices in serious gaming rehabilitation systems, including encephalogram [142-144] to monitor brain activity and ultrasound to estimate finger force [145]. A summary of the devices and their different combinations used to date is presented in Figure 4.

Figure 4. Overview of the devices used in the included studies. The green nodes indicate the sensors included in the *commercial sensor* category, while the orange nodes are those included in the *hardware development* category. The dark green nodes represent sensors that have been used individually or in combination with one of the sensors represented by the light green nodes.
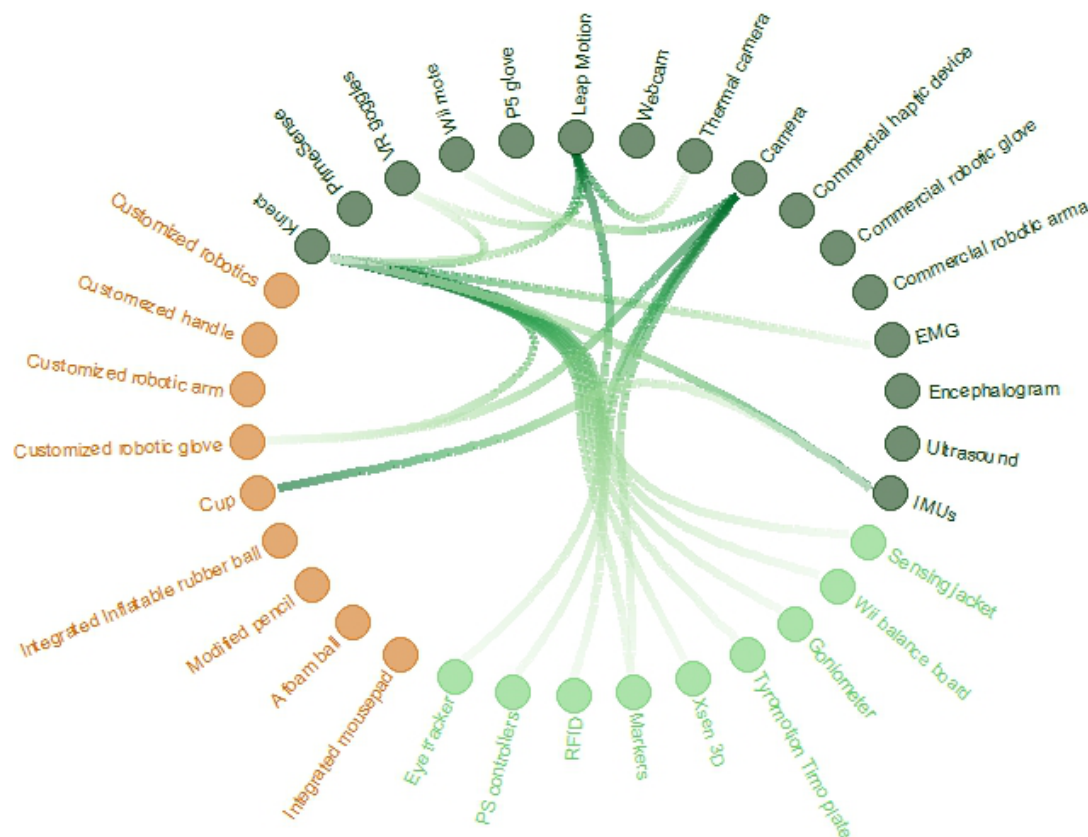
Hardware Development

Several different ideas of *hardware development* have been put forward in the developed devices. There are simple approaches that included objects such as a cup [54], combinations of objects and sensors such as a cup with IMU sensors [57], custom-made devices representing daily movements for elderly patients [146], a mousepad integrated with a CD motor [147], a foam ball and a modified pencil [148], and an inflatable rubber ball with an air pressure–sensing device [149]. Additionally, there are more complicated and time-consuming approaches identified in this category, including developments of exoskeleton robotic devices that focus on specific parts of the hand. These are studies in which robotic gloves were developed to control the fingers [35,42,61,79,150-156], and in which robotic arms were created to cover the surface from the shoulder to the wrist [4-11,48,49,110,139,157-161]. Additionally, many research teams have focused on the development of a handle to obtain control of the arm force and movement [12,56,135,162-173] or similar robotics [174-177]. A summary of these devices is shown in Figure 4.

Hardware Use Limitations

Despite the use of advanced technology, many studies that used commercial sensor devices mentioned limitations regarding the used hardware. For example, various studies [41,62,65,69,80,82] mentioned the possibility of Kinect's dysfunctionality in detecting movements or parts of the body. Moreover, several studies mentioned poor body part detection using different commercial sensors, such as poor hand detection from the Leap Motion device [103], poor detection with use of a camera [28], and poor detection in the combined use of the Leap Motion sensor with Oculus Rift VR goggles [104].

Limitations were also mentioned with respect to the *hardware development* category, including the need for enhanced calibration or upgraded components for better monitoring of accurate data. One study reported the need to improve a control strategy [139], while others mentioned general hardware issues [61,109,134,156,166]. In addition, some studies referred to the need for adjustments regarding the range of motion of the users [49,160] or the size of the hand [5]. Finally, one study pointed out difficulties in the use of the hardware system due to the poor design (ie, it was difficult for the user to put on the robotic glove and thus use it) [140].

Supervision Level

In general, the *supervision level* involved with each system was not always evident. Approximately half of the articles did not mention the home or clinical use of their systems, while some mentioned various supervision levels. It is worth noting that only 42.6% (72/169) of the included studies mentioned any supervision level for their proposed system; 3% (5/18) of the unsupervised cases belonged to the *hardware development*

category and 8%(13/18) belonged to the *commercial sensors* category.

System Testing With Users: Pilots and Trials

All of the included studies described a system that has been developed by the corresponding research team. Some of them included small pilot or limited-range trials, while others were complemented by subsequent studies reporting the results of pilot surveys. It is notable that the most frequent limitation in cases with no pilot study was the absence of clinical trials and the deficient testing of the system. Overall, 33.7%(57/169) of the studies tested their system with healthy subjects and therapists, whereas 27.1%(46/169) did not test the system at all. Although the remaining one-third of the research teams conducted clinical trials with patients, the majority of them included a very small number of patients, and also included healthy subjects in some cases to enlarge the sample. Among the 169 papers included in this scoping review, there were 2291 participants in the pilot and clinical trials. However, from the total number of participants, 689 were patients and 1602 were healthy volunteers, clinicians, therapists, and researchers.

System Evaluation

The *system evaluation* methods could not be easily categorized. This is because the research teams chose vastly different approaches for evaluation of their proposed system based on the target of the study and the available means at the time of publication. In 11.8% (20/169) of the included studies, no means of evaluation were mentioned regarding the introduced system, whereas 23.7% (40/169) of the studies seemed to focus only on users' or therapists' feedback about user experience via questionnaires and interviews. In these attempts, therapists and clinicians were given the opportunity to try the system with respect to the rehabilitation goals, the game's environment, and devices' safe use before being tested by patients.

Furthermore, 29.6%(50/169) of the studies used metrics regarding functional recovery via standard tests such as FMA, ARAT, and MAUULF; scores such as ROM for the elbow and forearm; and the Jamar strength test for strength of the hands. By using standard tests and scores as evaluation methods, researchers can measure the progress of a patient regarding motion dysfunction before and after the rehabilitation scheme. In addition, several studies tried to extract game features to introduce quantitative measurements for evaluation of patient improvement. Among these studies, 65%(49/75) did not analyze the evaluation methods, referring to them more generally as "data analysis" and providing descriptive statistics or as "monitoring data" in which the sessions were recorded using several sensors. In the next section, we discuss an extended analysis regarding the *extracted game features* of the studies.

Extracted Game Features

Quantitative measurements of the treatment outcome are critical for clinical rehabilitation practice, which constitute an objective method for evaluating the patient's medical progress. With these measurements, physicians can closely monitor the therapy process and adjust the treatment protocol individually. Among the studies included in this review, only a few described an assessment process of the patient's recovery status based on extracted features.

Some of these studies [21,65,98,124,126,128,137,147,155,164] used the time category *game target*, which was used to define metrics. These metrics were mainly classified into categories of hand movements or the duration, task, session completion of gestures, and reaction time.

Game performance was another consistent feature among the studies. In some cases, performance was associated with the score, and in other cases it was associated with task completion of the game target. Several studies [21,73,98,101,126,155] collated the score of the extracted features with standard clinical metrics (eg, box-and-block test and FMA) and suggested a strong correlation between them. By contrast, one study [98] reported that game achievements (the score regarding the number of coins collected) are not always an objective indicator of a patient's therapy progress. In addition, the task completion extracted features for a group of studies is the result of a patient's performance compared with a gold-standard method, which is usually an ideal movement trajectory (perhaps executed from a healthy subject) that the patient should follow, or an arithmetic measurement calculated after quantification of a specific hand movement [128,132,164,174]. Deviations from the gold standard are calculated and constitute the extracted features. For example, Lioulemes et al [132] first classified hand trajectories from patients with a support vector machine classifier and a hidden Markov model, and then calculated their deviations from the optimal trajectory as errors in space and time. For the second part of their analysis, they used dynamic time warping, which is a method for aligning optimally time-dependent sequences.

It is worth noting that several studies monitored other kinematic indicators of the patient's health condition that are not included in the *game target* classes for describing the patient's overall improvement [21,68,124,126,128,135,137,164]. Specifically, five studies [68,124,128,137,164] referred to the smoothness of the hand movement, or hand steadiness (jerk), during therapy sessions as the main feature. This jerk behavior is mainly described as abrupt changes in the direction of the hand's motion, and the way it is calculated may slightly differ from one study to another regarding the mathematical procedure employed. Furthermore, two studies [43,68] mentioned the use of hand trajectory curvature as a feature for kinematic analysis. In one study, the curvature of the hand trajectory was calculated as its deviation from an ideal straight line [178], whereas the second study computed the logarithm of the median of path curvature [179] to quantify "motion irregularity." Both studies included trajectory curvature in their criteria for measuring the arm's coordination. In addition, three studies [5,124,126] focused on fractionation as a game feature. Specifically, fractionation describes the ability to isolate the movement of the fingers and volitionally activate the motor units of the hand. Finally, two studies [136,137] focused on the muscle activation and caption of functional movement.

ROM and data regarding the angles of the hand during its motion constitute another significant group of features that have been commonly used by researchers and health professionals to

quantify therapy progress. Several studies [12,20,69,81,98,103,152] referred to the calculation of these kinds of features. Four studies [20,69,98,152] reported that they monitored ROM data (minimum, maximum, and average) for each single joint or exercise movement, and only one study [12] calculated ROM as a summative score of multiple movements or the difference. This extracted feature comprises a valuable tool, as it can be compared across sessions, subjects, or between the impaired and nonimpaired limb of each subject. According to one research team [98], plots of features regarding ROM facilitate the detection of distraction or movement pain during a patient's therapy session, thus resulting in a more effective performance diagnosis. Table 2 presents a summary of the *extracted game features.*

Table 2. Summary of the extracted game features.

Categories, Features	Number of studies
Time	
Time-related	10
Game performance	
Score/task completion	6
Golden standard comparison	4
Kinematic indicators	
Hand jerk	5
Trajectory curvature	2
Fractionation	2
Muscle activation	1
Range of motion	
Range of motion	7

It should be noted that this scoping review does not report every metric for each study, as our purpose was not to elaborate on how each study implemented the assessment of the patient's improvement but rather to outline and categorize the features extracted from the motion analysis process, excluding metrics of patients' engagement and motivation. Furthermore, in this attempt of feature extraction categorization, no distinction was made between studies that evaluated these features with a group of patients and those that conducted trials with healthy subjects.

Discussion

Principal Findings

With this scoping review, we aimed to explore the trends associated with deploying technologies for functional rehabilitation of the upper extremities. The results indicate that there has been increasing interest in these applications over time. The rapid evolution of technology contributes to new approaches concerning clinical practice and personalization of therapies. There is currently a wide range of sensors available for capturing motion (eg, Kinect, Leap Motion, IMUs) along with attempts to translate these technologies into an environment projected on computer screens, in VR headsets, or in AR image processing. The capabilities of various sensors related to motion, in conjunction with serious gaming, are used by research teams to develop contemporary systems for both doctors and patients.

One of the main advantages of this study is the overview of the current state of the field of upper extremity rehabilitation using serious games. As part of our research interest, we tried to investigate this topic to better understand the various approaches used to date. Based on this summary, we present a set of characteristics that depict a common direction and provide a complete picture of the sensors and technologies utilized to achieve the therapy purpose in terms of standard clinical practice.

The results indicate increasing research interest in the domain of serious gaming deployment in upper extremity rehabilitation. Based on the descriptive analysis, we can examine different aspects of this field of research. In particular, stroke seems to be a common *medical condition* for many research teams to trigger a study about upper limb mobility. This is understandable considering that stroke is the third leading cause of disability worldwide [180].

With respect to the factor *upper extremity part*, we found only one study published in 2016 by Hung et al [134] that presented a home-based rehabilitation system focused on the hand muscle. In addition, 26 of 36 studies that included the fingers as the hand part of focus used *commercial sensors* for their proposed rehabilitation system.

While reviewing the surveys included in our literature search, it is a safe assumption that computer graphic development in the last few decades has led to generalization of the term "VR." In 1992, Coates [181] defined VR as follows: "electronic simulations of environments experienced via head-mounted eye goggles and wired clothing, enabling the end-user to interact in realistic three-dimensional situations." Most of the studies included in this scoping review referred to their systems as "VR systems," but did not consider the original definition quoted above. These studies instead addressed a more generalized notion of VR that includes all systems that can simulate the real world (via 3D virtual environments generated by computer

graphics) and use sensors—but not necessarily VR headsets—for their interaction framework with the users. Besides VR systems, we found three studies that referred to their *game type* as "mixed reality" [42,62,89], which is a tabletop AR platform mixing VR, AR, and electronic board games for the proposed game.

With respect to the *game target*, every choice of game scenario was related to different aspects of rehabilitation schema. The game scenario combines the targeted upper extremity part, the moves that the patient needs to repeat for training, and the researcher's goal for an outcome that entertains and motivates the user. Increases in task completion and score are the most common scenarios differing in the content of the task based on the rehabilitation scheme.

Additionally, there is a broad selection of *sensors* and their combinations, as presented in Figure 4, that can be used for the creation of various rehabilitation systems. The most commonly combined sensors are Kinect, Leap Motion, and a camera, which seemed to be the most preferred devices over time, as described in further detail below. Technological development has provided researchers with more and more tools for testing their ideas, leading to new efficient implementations. Besides these new sensors, during the past decade, smartphones and tablets have entered our daily lives, and have rapidly become an integral part of life. Since 2011, researchers have been testing their capabilities in conjunction with upper extremity rehabilitation. It is worth noting that the wrist and fingers were the most commonly targeted *upper extremity parts* for rehabilitation using mobile or tablet app–based games, mainly because of the touchscreens.

In this study, we classified the used devices according to the *commercial sensors* and *hardware development* categories. Comparing the two *device development* categories, both showed differences in the proposed implementations. The larger number of publications related to the *commercialsensors* compared to the *hardware development* category implies greater interest in growing an idea of a game based on an advertised device. This higher interest may occur because it is more time-consuming or expensive to develop a new device than to explore the applications of already existing brands. Nevertheless, in the *commercial sensors* category, some studies reported poor hand part detection during sensor use, as elaborated upon in the Hardware Use Limitations subsection, deploying problems in practice. In addition, in the *hardware development* category, researchers have used the opportunity to develop a device based on the targeted hand part; however, depending on the case, a customized device could raise problems such as difficulties in use and adjustments for every hand size. Furthermore, many of the rehabilitation schemes included in the *commercial sensors* category targeted a home-based system using portable devices. However, for the *hardware development* category, the fragile and limited customized devices require supervised use, which poses a challenge for home-based trials.

Figure 5 presents a timeline of the devices reported in the literature over the past three decades. All of the categories presented in this figure include devices that may belong either to the *hardware development* or *commercial sensors* category. The aim of this figure is to present the use of every individual sensor in the research on upper limb rehabilitation over time. Since 1999, researchers have extensively studied glove sensors and robotics, a category including arms and handles. In addition, it is evident that since the first release of Kinect in 2010, there has been continuous interest in its use in upper extremity rehabilitation over the years. It is worth noting that until the release of Kinect, several studies were using the Wii remote, with the majority published in 2011, whereas after this point, there were only a few such attempts reported in 2013 and 2016. By contrast, Kinect gained increasing interest from 2011 to 2019. The Leap Motion sensor was first released in 2010, but the first attempts to use it in upper limb rehabilitation were only reported 4 years later in 2014, and the highest number of papers published in this field appeared 3 years later in 2017. Moreover, an inverse relation was observed between studies published on the Kinect sensor and the use of cameras over the years. Finally, the category biosignals includes biosensors such as EMG and electroencephalogram. Since 2011, these sensors have been consistently used in many studies (Figure 5).

Figure 5. Summary of devices reported in the literature over time. Groups of sensors used for studies of upper limb rehabilitation are presented in a timeline. The groups include devices from both the *hardware development* and *commercial sensors* categories.
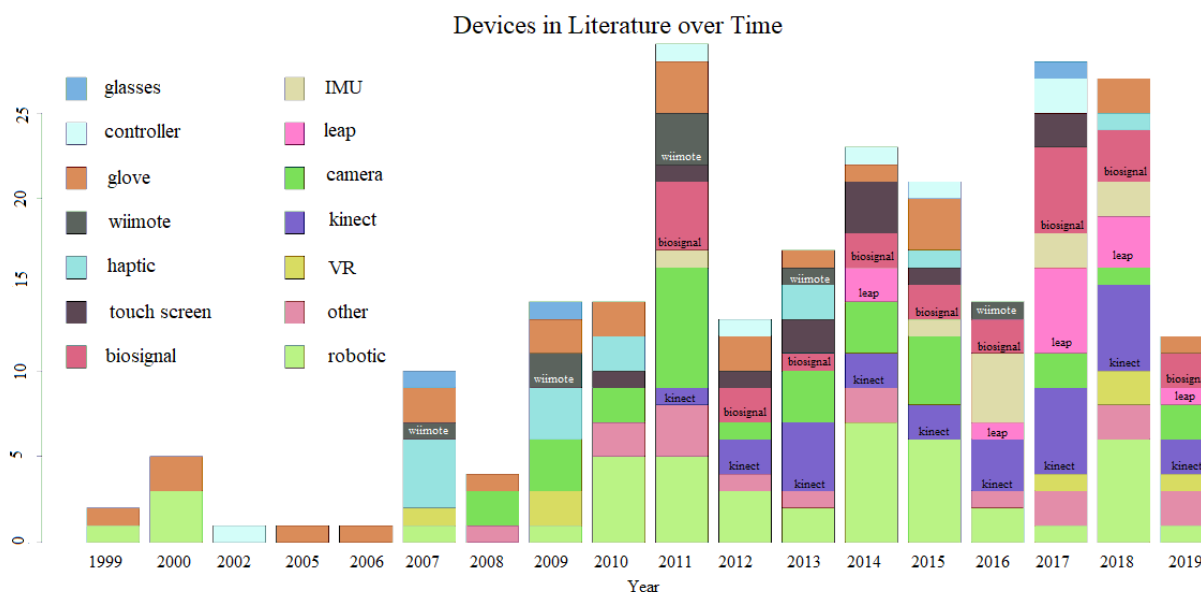
Based on our results, one-third of the introduced systems conducted clinical trials with patients to test their implementation. Although one-third of the included studies constitutes a sufficient number of attempts, in most of these cases, there is a need for a higher number of subjects and better testing of the rehabilitation scheme efficiency over time.

Concerning the *evaluation* of the proposed systems, many studies did not include any reference about the evaluation process of the patient's health or have not provided sufficient evidence about the assessment of the system. This is probably because most of this research was conducted at an early stage of development, prior to any related clinical trials, or because the researchers only aimed to introduce an idea about a rehabilitation system, mentioning their observations related to technical aspects as a secondary aim. Nevertheless, system evaluation of the tested rehabilitation schemes has typically been conducted using questionnaires or interviews, individually or in combination, about the game and the experience in general, medical standard tests examining improvement of the motion, and extracted game features about motion analysis or game performance. Notably, 68.7% (101/147) of the studies that included some form of assessment for the introduced rehabilitation game mentioned, among others, user experience as a factor for the system's estimation. In addition to user experience assessment, the most common system evaluation method involving patients was the use of standard clinical tests. Efforts for the creation of quantitative measurements of game-based treatment constitutes an attempt to provide evidence about the efficiency of the rehabilitation scheme and to personalize clinical practice.

Game features such as visual feedback of user actions and reward mechanisms via score/points or goal achievement were present in all rehabilitation gamification attempts. These features, which are an essential part of one of the biggest industries of the present day (ie, the game industry), are known to induce user engagement and are a core part of the rationale behind the gamification of health treatment protocols. In this

regard, an increase in patients' interest as a motive for investing in gamified approaches was taken for granted in most cases, although some studies also provided results from questionnaires that confirmed the above assumption.

Limitations

The broadness of the field of upper limb rehabilitation using serious games constitutes a limitation leading to many potentially included studies for this review. There was a significant number of studies, each suggesting different ways to approach the rehabilitation scheme but with poor sources or minimum attempts. Many conference papers have been published over the years introducing thoughts and preliminary results, but with no further analysis and implementation of their idea for rehabilitation. Although our exclusion criteria limited the range of the existing literature to some extent, this review includes several uncompleted attempts. Moreover, since we used specific keywords such as "upper-extremity," "rehabilitation," and "serious game" in different combinations, in an attempt to focus on the area of interest, we concede that some surveys in the field may have been excluded. Nevertheless, we are confident that the remaining studies that met all of the inclusion criteria can reflect the state of the field of upper extremity rehabilitation employing serious games, thereby assuring the reliability of our conclusions.

In addition, a limitation of this study is the lack of categorization based on gross motor vs fine motor or testing usability vs testing effectiveness. The many differently structured papers in combination with the heterogeneity in the provided information made such categorization very complicated and led us to the decision not to include these categories. It may be worth analyzing these categories separately to obtain an overview of this field in a different study.

Finally, due to the rapid technological progress, we consider another limitation to be the fact that this review includes studies only published up to June 2019.

Conclusion

Upper extremity motor dysfunction is a common problem that requires rehabilitation. Researchers studying the engagement of patients to the rehabilitation schemes have established several ways to develop more amusing training sets to better motivate patients. Technological progress constitutes an ally of these attempts, allowing for the combination of a traditional rehabilitation routine with serious games. In the last two decades, there has been a significant number of publications regarding upper limb rehabilitation using serious games, which is a field that continues to evolve based on user experience. Our goal regarding this review was to provide a complete overview of the field based on published studies over the years. Overall, this scoping review highlights several facts that point to the usefulness of serious games in rehabilitation in future medical procedures, as well as several weaknesses and challenges that have to be addressed. Despite the numerous attempts for establishing and evaluating game-based rehabilitation systems, more evidence is needed considering such systems not only as a means for patient motivation but also as an actual means for achieving upper extremity functionality improvement. In this vein, despite the challenges in the generalization and comparability of specific game decisions and implementations, it is important to support the efforts for the creation of quantitative measurements of game-based treatment, performance and outcome, and build evidence of its clinical value. In this direction, it would be important to work toward creating a framework for the therapeutic use of such gamified approaches, including the optimal dosage, personalization means, adaptations over time, session performance assessment, and therapeutic outcome. Such a therapeutic framework could enable the synthesis of more solid clinical evidence around game-based treatment, and eventually its incorporation in the clinical routine.

Authors' Contributions

IC, EK, and IL conceived of the study. All authors participated in the design of the study and drafted the manuscript. Data management was conducted by EK, IL, and DF, and advised by IC. All authors edited the manuscript, and read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Fotopoulos D, Kilintzis V, Chytas A, Mavromoustakos P, Loizidis T, Chouvarda I. Gamifying Motion Control Assessments Using Leap Motion Controller. Stud Health Technol Inform 2018;251:63-66. [Medline: 29968602]
2. Chytas A, Fotopoulos D, Kilintzis V, Loizidis T, Chouvarda I. Upper limb movement analysis of patients with neuromuscular disorders using data from a novel rehabilitation gaming platform. 2019 Presented at: XV Mediterranean Conference on Medical and Biological Engineering and Computing – MEDICON; September 26-28, 2019; Coimbra, Portugal p. 661-668. [doi: 10.1007/978-3-030-31635-8_79]
3. Milord: Modern Interface Platform for Motor Control and Learning on People with Motor Disorders. URL: http://milord.med.auth.gr/node/8 [accessed 2020-11-21]
4. Popescu V, Burdea G, Bouzit M, Girone M, Hentz V. PC-based telerehabilitation system with force feedback. Stud Health Technol Inform 1999;62:261-267. [Medline: 10538369]
5. Jack D, Boian R, Merians A. A virtual reality-based exercise program for stroke rehabilitation. 2000 Presented at: Annual ACM Conference on Assistive Technologies; November 13-15, 2000; Arlington, Virginia p. 56-63. [doi: 10.1145/354324.354340]
6. Popescu VG, Burdea GC, Bouzit M, Hentz VR. A virtual-reality-based telerehabilitation system with force feedback. IEEE Trans Inf Technol Biomed 2000 Mar;4(1):45-51. [doi: 10.1109/4233.826858] [Medline: 10761773]
7. Jack D, Boian R, Merians AS, Tremaine M, Burdea GC, Adamovich SV, et al. Virtual reality-enhanced stroke rehabilitation. IEEE Trans Neural Syst Rehabil Eng 2001 Sep;9(3):308-318. [doi: 10.1109/7333.948460] [Medline: 11561668]
8. Burdea GC, Cioi D, Martin J, Fensterheim D, Holenski M. The Rutgers Arm II rehabilitation system--a feasibility study. IEEE Trans Neural Syst Rehabil Eng 2010 Oct;18(5):505-514. [doi: 10.1109/TNSRE.2010.2052128] [Medline: 20529757]
9. Rabin B, Burdea G, Roll D, Hundal J, Damiani F, Pollack S. Integrative rehabilitation of elderly stroke survivors: the design and evaluation of the BrightArm™. Disabil Rehabil Assist Technol 2012 Jul;7(4):323-335 [FREE Full text] [doi: 10.3109/17483107.2011.629329] [Medline: 22107353]

10.   House G, Burdea G, Polistico K, Roll D, Kim J, Grampurohit N, et al. Integrative rehabilitation of residents chronic post-stroke in skilled nursing facilities: the design and evaluation of the BrightArm Duo. Disabil Rehabil Assist Technol 2016 Nov 28;11(8):683-694. [doi: 10.3109/17483107.2015.1068384] [Medline: 26218427]

11.   House G, Burdea G, Polistico K, Roll D, Kim J, Damiani F, et al. BrightArm duo integrative rehabilitation for post-stroke maintenance in skilled nursing facilities. : IEEE; 2015 Presented at: InInternational Conference on Virtual Rehabilitation (ICVR); June 9-12, 2015; Valencia p. 207-214. [doi: 10.1109/icvr.2015.7358594]

12.   Jarus T, Shavit S, Ratzon N. From hand twister to mind twister: computer-aided treatment in traumatic wrist fracture. Am J Occup Ther 2000;54(2):176-182. [doi: 10.5014/ajot.54.2.176] [Medline: 10732179]

13.   Reinkensmeyer DJ, Pang CT, Nessler JA, Painter CC. Web-based telerehabilitation for the upper extremity after stroke. IEEE Trans Neural Syst Rehabil Eng 2002 Jun;10(2):102-108. [doi: 10.1109/TNSRE.2002.1031978] [Medline: 12236447]

14.   Meijer HA, Graafland M, Goslings JC, Schijven MP. Systematic Review on the Effects of Serious Games and Wearable Technology Used in Rehabilitation of Patients With Traumatic Bone and Soft Tissue Injuries. Arch Phys Med Rehabil 2018 Sep;99(9):1890-1899. [doi: 10.1016/j.apmr.2017.10.018] [Medline: 29138050]

15.   Ang KK, Guan C. Brain-Computer Interface in Stroke Rehabilitation. J Comput Sci Eng 2013 Jun 30;7(2):139-146. [doi: 10.5626/jcse.2013.7.2.139]

16.   Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]

17.   Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Method 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]

18.   Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]

19.   Oña ED, Balaguer C, Cano-de la Cuerda R, Collado-Vázquez S, Jardón A. Effectiveness of Serious Games for Leap Motion on the Functionality of the Upper Limb in Parkinson's Disease: A Feasibility Study. Comput Intell Neurosci 2018;2018:7148427. [doi: 10.1155/2018/7148427] [Medline: 29849550]

20.   Schneider JC, Ozsecen MY, Muraoka NK, Mancinelli C, Della Croce U, Ryan CM, et al. Feasibility of an Exoskeleton-Based Interactive Video Game System for Upper Extremity Burn Contractures. PM & R 2016 May;8(5):445-452. [doi: 10.1016/j.pmrj.2015.09.012] [Medline: 26409198]

21.   Ustinova KI, Leonard WA, Cassavaugh ND, Ingersoll CD. Development of a 3D immersive videogame to improve arm-postural coordination in patients with TBI. J Neuroeng Rehabil 2011 Oct 31;8:61 [FREE Full text] [doi: 10.1186/1743-0003-8-61] [Medline: 22040301]

22.   Covarrubias M, Bordegoni M, Cugini U. Haptic Trajectories for Assisting Patients during Rehabilitation of Upper Extremities. Comput Aid Design Appl 2014 Oct 09;12(2):218-225. [doi: 10.1080/16864360.2014.962434]

23.   Ocampo R, Tavakoli M. Visual-haptic colocation in robotic rehabilitation exercises using a 2D augmented-reality display. 2019 Presented at: InInternational Symposium on Medical Robotics (ISMR). IEEE; April 3-5, 2019; Atlanta, GA. [doi: 10.1109/ismr.2019.8710185]

24.   Esfahlani S, Wilson G. Development of rehabilitation system (RehabGame) through Monte-Carlo tree search algorithm using kinect and Myo sensor interface. 2017 Presented at: 2017 Computing Conference; July 18-20, 2017; London, UK. [doi: 10.1109/sai.2017.8252217]

25.   Tang H, Feng Z, Xu T, Yang X. VR system for active hand rehabilitation training. 2017 Jul Presented at: Ninth International Conference on Information, Cybernetics and Computational Social Systems (ICCSS). IEEE; July 24-26, 2017; Dalian, China. [doi: 10.1109/iccss.2017.8091432]

26.   Arif A, Ulfa MN, Dharma D, Rizky AM, Prabowo T. An interactive Kinect-based game development for shoulder injury rehabilitation. 2018 Presented at: 2018 5th International Conference on Data and Software Engineering (ICoDSE); November 7-8, 2018; Mataram, Lombok, Indonesia. [doi: 10.1109/icodse.2018.8705844]

27.   Karime A, Mahfujur A, Saddik A, Gueaieb W. RehaBall: Rehabilitation of upper limbs with a sensory-integrated stress ball. 2011 Presented at: 2011 IEEE International Workshop on Haptic Audio Visual Environments and Games; October 14-17, 2011; Hebei, China. [doi: 10.1109/have.2011.6088388]

28.   Burke J, Morrow P, McNeill M, McDonough S, Charles D. Vision based games for upper-limb stroke rehabilitation. In: International Machine Vision and Image Processing Conference.: IEEE; 2008 Presented at: 2008 International Machine Vision and Image Processing Conference; September 3-5, 2008; Portrush, Ireland. [doi: 10.1109/imvip.2008.16]

29.   King M, Hale L, Pekkari A, Persson M, Gregorsson M, Nilsson M. An affordable, computerised, table-based exercise system for stroke survivors. Disabil Rehabil Assist Technol 2010 Jul;5(4):288-293. [doi: 10.3109/17483101003718161] [Medline: 20302419]

30.   Aung Y, Al-jumaily A. Augmented reality based reaching exercise for shoulder rehabilitation. 2011 Presented at: i-CREATe '11: Proceedings of the 5th International Conference on Rehabilitation Engineering & Assistive Technology; July 2011; Bangkok, Thailand. [doi: 10.5555/2500753.2500812]

31. Aung Y, Al-jumaily A. Development of augmented reality rehabilitation games integrated with biofeedback for upper limb. 2011 Presented at: i-CREATe '11: Proceedings of the 5th International Conference on Rehabilitation Engineering & Assistive Technology; July 2011; Bangkok, Thailand. [doi: 10.1145/1328491.1328504]

32. Viriyasaksathian B, Khemmachotikun S, Kaimuk P, Wongsawat Y. EMG-based upper-limb rehabilitation via music synchronization with augmented reality. 2011 Presented at: 2011 IEEE International Conference on Robotics and Biomimetics; December 7-11, 2011; Karon Beach, Phuket, Thailand. [doi: 10.1109/robio.2011.6181738]

33. Regenbrecht H, McGregor G, Ott C, Hoermann S, Schubert T, Hoermann J, et al. Out of reach? -- A novel AR interface approach for motor rehabilitation. 2011 Presented at: 2011 10th IEEE International Symposium on Mixed and Augmented Reality; October 26-29, 2011; Basel, Switzerland. [doi: 10.1109/ismar.2011.6162891]

34. Aung Y, Al-Jumaily A. Rehabilitation exercise with real-time muscle simulation based EMG and AR. 2011 Presented at: 2011 11th International Conference on Hybrid Intelligent Systems (HIS); December 5-8, 2011; Melacca, Malaysia. [doi: 10.1109/his.2011.6122181]

35. Choi Y. Ubi-REHAB: An android-based portable Augmented Reality stroke rehabilitation system using the eGlove for multiple participants. 2011 Presented at: 2011 International Conference on Virtual Rehabilitation; June 27-29, 2011; Zurich, Switzerland. [doi: 10.1109/icvr.2011.5971818]

36. Regenbrecht H, Hoermann S, McGregor G, Dixon B, Franz E, Ott C, et al. Visual manipulations for motor rehabilitation. Comput Graph 2012 Nov;36(7):819-834. [doi: 10.1016/j.cag.2012.04.012]

37. Leon N, Bhatt S, Al-Jumaily A. Augmented reality game based multi-usage rehabilitation therapist for stroke patients. Int J Smart Sens Intel Syst 2014;7(3):1044-1058. [doi: 10.21307/ijssis-2017-693]

38. Hoermann S, Collins J, Regenbrecht H. Learning from rehabilitation: A bi-manual interface approach. 2015 Presented at: 2015 IEEE Symposium on 3D User Interfaces (3DUI); March 23-24, 2015; Arles, France. [doi: 10.1109/3dui.2015.7131751]

39. Ying W, Aimin W. Augmented reality based upper limb rehabilitation system. 2017 Presented at: 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI); October 20-22, 2017; Yangzhou, China. [doi: 10.1109/icemi.2017.8265843]

40. Xue Y, Zhao L, Xue M, Fu J. Gesture interaction and augmented reality based hand rehabilitation supplementary System. 2018 Presented at: 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC); October 12-14, 2018; Chongqing, China. [doi: 10.1109/iaeac.2018.8577476]

41. Melero M, Hou A, Cheng E, Tayade A, Lee SC, Unberath M, et al. Upbeat: Augmented Reality-Guided Dancing for Prosthetic Rehabilitation of Upper Limb Amputees. J Healthc Eng 2019;2019:2163705. [doi: 10.1155/2019/2163705] [Medline: 31015903]

42. Halic T, Kockara S, Demirel D, Willey M, Eichelberger K. MoMiReS: Mobile mixed reality system for physical & occupational therapies for hand and wrist ailments. 2014 Presented at: 2014 IEEE Innovations in Technology Conference; May 16, 2014; Warwick, RI. [doi: 10.1109/innotek.2014.6877376]

43. Lourenco F, Postolache O, Postolache G. Tailored virtual reality and mobile application for motor rehabilitation. 2018 Presented at: 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC); May 14-17, 2018; Houston, TX. [doi: 10.1109/i2mtc.2018.8409572]

44. Zirbel C, Zhang X, Hughes C. The VRehab System: A Low-Cost Mobile Virtual Reality System for Post-Stroke Upper Limb Rehabilitation for Medically Underserved Populations. 2018 Presented at: 2018 IEEE Global Humanitarian Technology Conference (GHTC); October 18-21, 2018; San Jose, CA. [doi: 10.1109/ghtc.2018.8601885]

45. Octavia JR, Natasha L. Design of a Mobile Game Application to Support Hand Rehabilitation of Stroke Patients in Indonesia. IJTech 2017 Apr 29;8(2):250. [doi: 10.14716/ijtech.v8i2.6167]

46. Guo J, Smith T, Messing D, Tang Z, Lawson S, Feng J. ARMStrokes: A Mobile App for Everyday Stroke Rehabilitation. 2015 Presented at: ASSETS '15: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility; October 26-28, 2015; Lisbon, Portugal. [doi: 10.1145/2700648.2811337]

47. Megard C, Bouchigny S, Martin A, Goulamhoussen R, Bartholier L, Foulon P, et al. Ergotact: Including Force-based Activities into Post-stroke Rehabilitation. New York, NY: ACM Press; 2019 Presented at: CHI EA '19: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems; May 2019; Glasgow, Scotland. [doi: 10.1145/3290607.3312920]

48. English BA, Howard AM. Encouraging specific intervention motions via a robotic system for rehabilitation of hand function: A healthy pilot study. 2014 Presented at: 2014 IEEE Symposium on Computational Intelligence in Robotic Rehabilitation and Assistive Technologies (CIR2AT); December 9-12, 2014; Orlando, FL. [doi: 10.1109/cirat.2014.7009741]

49. English BA, Howard AM. Engagement study of an integrated rehabilitation robotic tablet-based gaming system. 2014 Presented at: 2014 IEEE International Workshop on Advanced Robotics and its Social Impacts; September 11-13, 2014; Evanston, IL. [doi: 10.1109/arso.2014.7020973]

50. Boulanger C, Boulanger A, Greef L, Kearney A, Sobel K, Transue R, et al. Stroke rehabilitation with a sensing surface. 2013 Presented at: CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 27-May 2, 2013; Paris, France. [doi: 10.1145/2470654.2466160]

51. Carabeo CGG, Dalida CMM, Padilla EMZ, Rodrigo MMT. Stroke Patient Rehabilitation: A Pilot Study of an Android-Based Game. Simul Gaming 2014 Jun 24;45(2):151-166. [doi: 10.1177/1046878114531102]

XSL•FO

RenderX

52. Delbressine F, Timmermans A, Beursgens L, de Jong M, van Dam A, Verweij D, et al. Motivating arm-hand use for stroke patients by serious games. Annu Int Conf IEEE Eng Med Biol Soc 2012;2012:3564-3567. [doi: 10.1109/EMBC.2012.6346736] [Medline: 23366697]

53. Lee C, Chiu Y, Kao H, Chen I, Lee I, Ho W, et al. A Body-Sensed Motor Assessment System for Stroke Upper-Limb Rehabilitation: A Preliminary Study. 2013 Presented at: 2013 IEEE International Conference on Systems, Man, and Cybernetics; October 13-16, 2013; Manchester, UK. [doi: 10.1109/smc.2013.652]

54. Carneiro F, Tavares R, Rodrigues J, Abreu P, Restivo MT. A Gamified Approach for Hand Rehabilitation Device. Int J Onl Eng 2018 Jan 25;14(01):179. [doi: 10.3991/ijoe.v14i01.7793]

55. Jayasree-Krishnan V, Gamdha D, Goldberg B, Ghosh S, Raghavan P, Kapila V. A novel task-specific upper-extremity rehabilitation system with interactive game-based interface for stroke patients. 2019 Presented at: 2019 International Symposium on Medical Robotics (ISMR); April 3-5, 2019; Atlanta, GA. [doi: 10.1109/ismr.2019.8710184]

56. Unluhisarcikli O, Weinberg B, Sivak M, Mirelman A, Bonato P, Mavroidis C. A robotic hand rehabilitation system with interactive gaming using novel electro-rheological fluid based actuators. 2010 Presented at: 2010 IEEE International Conference on Robotics and Automation; May 3-7, 2010; Anchorage, AK. [doi: 10.1109/robot.2010.5509806]

57. Carneiro F, Santos B, Abreu P, Restivo M. A tool for grip evaluation and learning. 2016 Presented at: 2016 13th International Conference on Remote Engineering and Virtual Instrumentation (REV); February 24-26, 2016; Madrid, Spain. [doi: 10.1109/rev.2016.7444489]

58. Goršič M, Cikajlo I, Novak D. Competitive and cooperative arm rehabilitation games played by a patient and unimpaired person: effects on motivation and exercise intensity. J Neuroeng Rehabil 2017 Mar 23;14(1):23 [FREE Full text] [doi: 10.1186/s12984-017-0231-4] [Medline: 28330504]

59. Lafond I, Qiu Q, Adamovich SV. Design of a customized virtual reality simulation for retraining upper extremities after stroke. 2010 Presented at: Proceedings of the 2010 IEEE 36th Annual Northeast Bioengineering Conference (NEBEC); March 26-28, 2010; New York, NY. [doi: 10.1109/nebc.2010.5458130]

60. Li C, Rusák Z, Hou Y, Young C, Ji L. Upper limb motor rehabilitation integrated with video games focusing on training fingers' fine movements. Int J Robot Autom 2014;29(4):359-368. [doi: 10.2316/journal.206.2014.4.206-3984]

61. Alexandre R, Postolache O. Wearable and IoT technologies application for physical rehabilitation. 2018 Presented at: 2018 International Symposium in Sensing and Instrumentation in IoT Era (ISSI); September 6-7, 2018; Shanghai, China. [doi: 10.1109/issi.2018.8538058]

62. Robertson C, Vink L, Regenbrecht H, Lutteroth C, Wunsche B. Mixed reality Kinect Mirror box for stroke rehabilitation. 2013 Presented at: 2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013); November 27-29, 2013; Wellington, New Zealand. [doi: 10.1109/ivcnz.2013.6727021]

63. Cargnin DJ, Cordeiro d'Ornellas M, Cervi Prado AL. A Serious Game for Upper Limb Stroke Rehabilitation Using Biofeedback and Mirror-Neurons Based Training. Stud Health Technol Inform 2015;216:348-352. [Medline: 26262069]

64. d'Ornellas M, Cargnin D, Prado A. Thoroughly approach to upper limb rehabilitation using serious games for intensive group physical therapy or individual biofeedback training. 2014 Presented at: 2014 Brazilian Symposium on Computer Games and Digital Entertainment; November 12-14, 2014; Porto Alegre, Brazil. [doi: 10.1109/sbgames.2014.22]

65. Funabashi A, Aranha RV, Silva T, Monteiro C, Silva W, Nunes F. AGaR: A VR Serious Game to Support the Recovery of Post-Stroke Patients. 2017 Presented at: 2017 19th Symposium on Virtual and Augmented Reality (SVR); November 1-4, 2017; Curitiba, Brazil. [doi: 10.1109/svr.2017.15]

66. Shiratuddin M, Rai S, Krishnan G, Newton M, Wang X, Sohel F, et al. A usability evaluation of Neuromender's upper limb game-based rehabilitation system for stroke survivors. 2017 Presented at: 2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH); April 2-4, 2017; Perth, Australia. [doi: 10.1109/segah.2017.7939258]

67. Nixon M, Howard A. Applying gaming principles to virtual environments for upper extremity therapy games. 2013 Presented at: 2013 IEEE International Conference on Systems, Man, and Cybernetics; October 13-16, 2013; Manchester, UK. [doi: 10.1109/smc.2013.585]

68. Bai J, Song A, Li H. Design and Analysis of Cloud Upper Limb Rehabilitation System Based on Motion Tracking for Post-Stroke Patients. Appl Sci 2019 Apr 18;9(8):1620. [doi: 10.3390/app9081620]

69. Sen SL, Fai YC, Ming ESL, Xiang KK, Fai YC, Khan QI. Enhancing effectiveness of virtual reality rehabilitation system: Durian Runtuh. 2015 Presented at: 2015 10th Asian Control Conference (ASCC); May 31-June 3, 2015; Kota Kinabalu, Malaysia. [doi: 10.1109/ascc.2015.7244607]

70. Engsberg J, Behar C, Lustick M, Foreman M, Webb J. Personalized Virtual Reality for Upper Extremity Rehabilitation: Moving from the Clinic to a Home Exercise Program. Intellect Disabl Diagn J 2016 Oct 19;4(3):160-169. [doi: 10.6000/2292-2598.2016.04.03.3]

71. Valdés BA, Hilderman CGE, Hung C, Shirzad N, Van der Loos HFM. Usability testing of gaming and social media applications for stroke and cerebral palsy upper limb rehabilitation. Annu Int Conf IEEE Eng Med Biol Soc 2014;2014:3602-3605. [doi: 10.1109/EMBC.2014.6944402] [Medline: 25570770]

72. Ona E, Balaguer C, Jardon A. Towards a framework for rehabilitation and assessment of upper limb motor function based on Serious Games. 2018 Presented at: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH); May 16-18, 2018; Vienna, Austria. [doi: 10.1109/segah.2018.8401346]

XSL•FO
RenderX

73. Sookhanaphibarn K, Phukongchai W, Santad T, Choensawat W. Towards bilateral upper-limb rehabilitation after stroke using Kinect game. 2018 Presented at: 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE); October 9-12, 2018; Nara, Japan. [doi: 10.1109/gcce.2018.8574861]

74. Esfahlani S, Cirstea S, Sanaei A, Wilson G. An adaptive self-organizing fuzzy logic controller in a serious game for motor impairment rehabilitation. 2017 Presented at: 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE); June 19-21, 2017; Edinburgh, UK. [doi: 10.1109/isie.2017.8001435]

75. Saini S, Rambli D, Sulaiman S, Zakaria M, Mohd SS. A low-cost game framework for a home-based stroke rehabilitation system. 2012 Presented at: 2012 International Conference on Computer & Information Science (ICCIS); June 12-14, 2012; Kuala Lumpeu, Malaysia. [doi: 10.1109/iccisci.2012.6297212]

76. Trombetta M, Bazzanello Henrique PP, Brum MR, Colussi EL, De Marchi ACB, Rieder R. Motion Rehab AVE 3D: A VR-based exergame for post-stroke rehabilitation. Comput Methods Programs Biomed 2017 Nov;151:15-20. [doi: 10.1016/j.cmpb.2017.08.008] [Medline: 28946996]

77. Borghese NA, Pirovano M, Lanzi PL, Wüest S, de Bruin ED. Computational Intelligence and Game Design for Effective At-Home Stroke Rehabilitation. Games Health J 2013 Apr;2(2):81-88 [FREE Full text] [doi: 10.1089/g4h.2012.0073] [Medline: 24761321]

78. Pinto J, Carvalho H, Chambel G, Ramiro J, Goncalves A. Adaptive gameplay and difficulty adjustment in a gamified upper-limb rehabilitation. 2018 Presented at: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH); May 16-18, 2018; Vienna, Austria. [doi: 10.1109/segah.2018.8401363]

79. Huang M, Chen E, Xu W, Sarrafzadeh M. Gaming for upper extremities rehabilitation. 2011 Presented at: WH '11: Proceedings of the 2nd Conference on Wireless Health; October 2011; San Diego, California. [doi: 10.1145/2077546.2077576]

80. Viegas V, Postolache O, Pereira J, Girao P. NUI therapeutic serious games with metrics validation based on wearable devices. 2016 Presented at: 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings; May 23-26, 2016; Taipei, Taiwan. [doi: 10.1109/i2mtc.2016.7520550]

81. Tsoupikova D, Triandafilou K, Rupp G, Preuss F, Kamper D. Multi-user virtual reality therapy for post-stroke hand rehabilitation at home. J Syst Cybernet Informat 2016;14(2):67-71. [doi: 10.1145/2996376.2996387]

82. Ma M, Proffitt R, Skubic M. Validation of a Kinect V2 based rehabilitation game. PLoS One 2018 Aug 24;13(8):e0202338 [FREE Full text] [doi: 10.1371/journal.pone.0202338] [Medline: 30142631]

83. Sadihov D, Migge B, Gassert R, Kim Y. Prototype of a VR upper-limb rehabilitation system enhanced with motion-based tactile feedback. 2013 Presented at: 2013 World Haptics Conference (WHC); April 14-17, 2013; Daejeon, South Korea. [doi: 10.1109/whc.2013.6548450]

84. Bower KJ, Louie J, Landesrocha Y, Seedy P, Gorelik A, Bernhardt J. Clinical feasibility of interactive motion-controlled games for stroke rehabilitation. J Neuroeng Rehabil 2015 Aug 02;12:63 [FREE Full text] [doi: 10.1186/s12984-015-0057-x] [Medline: 26233677]

85. Shin J, Ryu H, Jang SH. A task-specific interactive game-based virtual reality rehabilitation system for patients with stroke: a usability test and two clinical experiments. J Neuroeng Rehabil 2014 Mar 06;11:32 [FREE Full text] [doi: 10.1186/1743-0003-11-32] [Medline: 24597650]

86. Dreifaldt U, Goude D, Rydmark M. A home-based Virtual Reality system for Stroke Rehabilitation. 2007 Presented at: 2007 Virtual Rehabilitation; September 27-29, 2007; Venice, Italy. [doi: 10.1109/icvr.2007.4362145]

87. Barzilay O, Wolf A. Adaptive rehabilitation games. J Electromyogr Kinesiol 2013 Feb;23(1):182-189. [doi: 10.1016/j.jelekin.2012.09.004] [Medline: 23141481]

88. Elor A, Teodorescu M, Kurniawan S. Project Star Catcher: A Novel Immersive Virtual Reality Experience for Upper Limb Rehabilitation. ACM Trans Access Comput 2018 Nov 22;11(4):1-25. [doi: 10.1145/3265755]

89. Ines D, Abdelkader G, Nadia H. Mixed reality serious games for post-stroke rehabilitation. In: UbiTheraGames. 2012 Presented at: Therapeutic Serious Games and Pervasive Computing; May 23, 2011; Dublin, Republic of Ireland. [doi: 10.4108/icst.pervasivehealth.2011.246102]

90. Lin J, Kelleher CL, Engsberg JR. Developing Home-Based Virtual Reality Therapy Interventions. Games Health J 2013 Feb;2(1):34-38. [doi: 10.1089/g4h.2012.0033] [Medline: 26196552]

91. Tsekleves E, Paraskevopoulos IT, Warland A, Kilbride C. Development and preliminary evaluation of a novel low cost VR-based upper limb stroke rehabilitation platform using Wii technology. Disabil Rehabil Assist Technol 2016;11(5):413-422. [doi: 10.3109/17483107.2014.981874] [Medline: 25391221]

92. Galego B, Simone L. Leveraging online virtual worlds for upper extremity Rehabilitation. 2007 Presented at: 2007 IEEE 33rd Annual Northeast Bioengineering Conference; March 10-11, 2007; Long Island, NY. [doi: 10.1109/nebc.2007.4413380]

93. Kaluarachchi C, Al-jumaily A. Self-rehabilitation based on user interactive environment. 2011 Presented at: i-CREATe '11: Proceedings of the 5th International Conference on Rehabilitation Engineering & Assistive Technology; July 2011; Bangkok, Thailand. [doi: 10.5555/2500753.2500814]

94. Burke J, McNeill M, Charles D, Morrow P, Crosbie J, McDonough S. Serious games for upper limb rehabilitation following stroke. 2009 Presented at: 2009 Conference in Games and Virtual Worlds for Serious Applications; March 23-24, 2009; Coventry, UK. [doi: 10.1109/vs-games.2009.17]

XSL•FO

RenderX

95. Kaluarachchi C, Aung Y, Al-Jumaily A. Virtual games based self rehabilitation for home therapy system. 2011 Presented at: 2011 11th International Conference on Hybrid Intelligent Systems (HIS); December 5-8, 2011; Melacca, Malaysia. [doi: 10.1109/his.2011.6122183]

96. Morrow K, Docan C, Burdea G, Merians A. Low-cost virtual rehabilitation of the hand for patients post-stroke. 2006 Presented at: 2006 International Workshop on Virtual Rehabilitation; August 29-30, 2006; New York, NY. [doi: 10.1109/iwvr.2006.1707518]

97. August K, Bleichenbacher D, Adamovich S. Virtual reality physical therapy: a telerehabilitation tool for hand and finger movement exercise monitoring and motor skills analysis. 2005 Presented at: Proceedings of the IEEE 31st Annual Northeast Bioengineering Conference, 2005; April 2-3, 2005; Hoboken, NJ. [doi: 10.1109/nebc.2005.1431931]

98. Afyouni I, Rehman FU, Qamar AM, Ghani S, Hussain SO, Sadiq B, et al. A therapy-driven gamification framework for hand rehabilitation. User Model User-Adap Inter 2017 Apr 24;27(2):215-265. [doi: 10.1007/s11257-017-9191-4]

99. Charles D, Pedlow K, McDonough S, Shek K, Charles T. Close range depth sensing cameras for virtual reality based hand rehabilitation. J Assist Technol 2014 Sep 09;8(3):138-149. [doi: 10.1108/jat-02-2014-0007]

100. Elnaggar A, Reichardt D. Digitizing the Hand Rehabilitation Using Serious Games Methodology with User-Centered Design Approach. 2016 Presented at: 2016 International Conference on Computational Science and Computational Intelligence (CSCI); December 15-17, 2016; Las Vegas, NV. [doi: 10.1109/csci.2016.0011]

101. Khademi M, Mousavi HH, McKenzie A, Dodakian L, Lopes C, Cramer S. Free-hand interaction with leap motion controller for stroke rehabilitation. 2014 Presented at: CHI EA '14: CHI '14 Extended Abstracts on Human Factors in Computing Systems; April 2014; Toronto, Ontario. [doi: 10.1145/2559206.2581203]

102. Alimanova M, Borambayeva S, Kozhamzharova D, Kurmangaiyeva N, Opsanova D, Tyulepberdinova G, et al. Gamification of Hand Rehabilitation Process Using Virtual Reality Tools: Using Leap Motion for Hand Rehabilitation. 2017 Presented at: 2017 First IEEE International Conference on Robotic Computing (IRC); April 10-12, 2017; Taichung, Taiwan. [doi: 10.1109/irc.2017.76]

103. Afyouni I, Qamar A, Hussain S, Ur RF, Sadiq B, Murad A. Motion-Based Serious Games for Hand Assistive Rehabilitation. 2017 Presented at: IUI '17 Companion: Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion; March 2017; Limassol, Cyprus. [doi: 10.1145/3030024.3040977]

104. Dias P, Silva R, Amorim P, Lains J, Roque E, Pereira ISF, et al. Using Virtual Reality to Increase Motivation in Poststroke Rehabilitation. IEEE Comput Grap Appl 2019 Jan;39(1):64-70. [doi: 10.1109/mcg.2018.2875630]

105. Postolache O, Lourenco F, Dias PJ, Girao P. Serious game for physical rehabilitation: Measuring the effectiveness of virtual and real training environments. 2017 Presented at: 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC); May 22-25, 2017; Turin, Italy. [doi: 10.1109/i2mtc.2017.7969978]

106. Yin Z, Xu H. A wearable rehabilitation game controller using IMU sensor. 2018 Presented at: 2018 IEEE International Conference on Applied System Invention (ICASI); April 13-17, 2018; Chiba, Japan. [doi: 10.1109/icasi.2018.8394459]

107. Han J, Lian S, Guo B, Li X, You A. Active rehabilitation training system for upper limb based on virtual reality. Adv Mech Eng 2017 Dec 08;9(12):168781401774338. [doi: 10.1177/1687814017743388]

108. Wittmann F, Lambercy O, Gonzenbach R, Raai M, Hover R, Held J, et al. Assessment-driven arm therapy at home using an IMU-based virtual reality system. 2015 Presented at: 2015 IEEE International Conference on Rehabilitation Robotics (ICORR); August 11-14, 2015; Singapore. [doi: 10.1109/icorr.2015.7281284]

109. Ribeiro T, Veloso A, Costa R. Conceptualization of PhysioFun game: A low-cost videogame for home-based stroke rehabilitation. 2016 Presented at: 2016 1st International Conference on Technology and Innovation in Sports, Health and Wellbeing (TISHW); December 1-3, 2016; Vila Real, Portugal. [doi: 10.1109/tishw.2016.7847787]

110. Tageldeen M, Elamvazuthi I, Perumal N, Ganesan T, Borboni A. Development of a 3D virtual platform for an upper arm robotic exoskeleton. 2016 Presented at: 2016 2nd IEEE International Symposium on Robotics and Manufacturing Automation (ROMA); September 25-27, 2016; Ipoh, Malaysia. [doi: 10.1109/roma.2016.7847819]

111. Prange G, Krabben T, Molier B, Kooij H, Jannink M. A low-tech virtual reality application for training of upper extremity motor function in neurorehabilitation. 2008 Presented at: 2008 Virtual Rehabilitation; August 25-27, 2008; Vancouver, BC. [doi: 10.1109/icvr.2008.4625113]

112. Aung Y, Al-Jumaily A, Anam K. A novel upper limb rehabilitation system with self-driven virtual arm illusion. Annu Int Conf IEEE Eng Med Biol Soc 2014;2014:3614-3617. [doi: 10.1109/EMBC.2014.6944405] [Medline: 25570773]

113. Chiang WT, Siew WC, Wai XL. Game-based human computer interaction using gesture recognition for rehabilitation. 2013 Presented at: 2013 IEEE International Conference on Control System, Computing and Engineering; November 29-December 1, 2013; Mindeb, Malaysia. [doi: 10.1109/iccsce.2013.6719987]

114. Burke JW, McNeill MDJ, Charles DK, Morrow PJ, Crosbie JH, McDonough SM. Optimising engagement for stroke rehabilitation using serious games. Vis Comput 2009 Aug 27;25(12):1085-1099. [doi: 10.1007/s00371-009-0387-4]

115. Crosbie JH, McNeill MDJ, Burke J, McDonough S. Utilising technology for rehabilitation of the upper limb following stroke: the Ulster experience. Phys Ther Rev 2013 Jul 19;14(5):336-347. [doi: 10.1179/108331909x12540993897892]

116. Pham N, Phan V, Nguyen V, Nguyen V. Game-based virtual rehabilitation system for upper extremity using low-cost camera. 2015 Presented at: 2015 8th Biomedical Engineering International Conference (BMEiCON); November 25-27, 2015; Pattaya, Thailand. [doi: 10.1109/bmeicon.2015.7399505]

117. Hung CT, Croft E, Van der Loos M. A wearable vibrotactile device for upper-limb bilateral motion training in stroke rehabilitation: A case study. 2015 Presented at: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); August 25-29, 2015; Milan, Italy. [doi: 10.1109/embc.2015.7319142]

118. Kim SJ, Han SY, Yang G, Kim J, Ahn B. Development of an interactive game-based mirror image hand rehabilitation system. Intel Serv Robotics 2019 Feb 20;12(2):149-157. [doi: 10.1007/s11370-018-00272-5]

119. Evett L, Burton A, Battersby S, Brown D, Sherkat N, Ford G, et al. Dual camera motion capture for serious games in stroke rehabilitation. 2011 Presented at: 2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH); November 16-18, 2011; Braga, Portugal. [doi: 10.1109/segah.2011.6165460]

120. Yang X, Wu X, Zhao Z, Li Y. Hand tele-rehabilitation in haptic virtual environment. 2007 Presented at: 2007 IEEE International Conference on Robotics and Biomimetics (ROBIO); December 15-18, 2007; Sanya, China. [doi: 10.1109/robio.2007.4522150]

121. Broeren J, Sunnerhagen KS, Rydmark M. Haptic virtual rehabilitation in stroke: transferring research into clinical practice. Phys Ther Rev 2013 Jul 19;14(5):322-335. [doi: 10.1179/108331909x12488667117212]

122. Auguste-Rampersad D, Medford G, Kanneh A. Haptics and post stroke rehabilitation of the hand: Immersive gaming using haptics to motivate patients to continue repetitive exercises needed for rehabilitation after stroke. 2018 Presented at: 2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE); April 28-29, 2018; Penang, Malaysia. [doi: 10.1109/iscaie.2018.8405510]

123. Xu Z, Yu H, Yan S. Motor rehabilitation training after stroke using haptic handwriting and games. 2010 Presented at: iCREATe '10: Proceedings of the 4th International Convention on Rehabilitation Engineering & Assistive Technology; July 21-24, 2010; Shanghai, China. [doi: 10.5555/1926058.1926089]

124. Adamovich S, Qiu Q, Talati B, Fluet G, Merians A. Design of a Virtual Reality-Based System For Hand and Arm Rehabilitation. 2007 Presented at: 2007 IEEE 10th International Conference on Rehabilitation Robotics; June 13-15, 2007; Noordwijk, Netherlands. [doi: 10.1109/icorr.2007.4428539]

125. Qinyin Q, Ramirez D, Saleh S, Adamovich S. NJIT-RAVR system for upper extremity rehabilitation in children with hemiplegia. 2009 Presented at: 2009 IEEE 35th Annual Northeast Bioengineering Conference; April 3-5, 2009; Boston, MA. [doi: 10.1109/nebc.2009.4967703]

126. Merians A, Lewis J, Qiu Q, Fluet G, Talati B, Adamovich S. Strategies for Incorporating Bilateral Training into a Virtual Environment. 2007 Presented at: 2007 IEEE/ICME International Conference on Complex Medical Engineering; May 23-27, 2007; Beijing, China. [doi: 10.1109/iccme.2007.4381948]

127. Qiu Q, Ramirez DA, Saleh S, Fluet GG, Parikh HD, Kelly D, et al. The New Jersey Institute of Technology Robot-Assisted Virtual Rehabilitation (NJIT-RAVR) system for children with cerebral palsy: a feasibility study. J Neuroeng Rehabil 2009 Nov 16;6:40 [FREE Full text] [doi: 10.1186/1743-0003-6-40] [Medline: 19917124]

128. Pareek S, Chembrammel P, Kesavadas T. Development and evaluation of haptics-based rehabilitation system. 2018 Presented at: 2018 International Symposium on Medical Robotics (ISMR); March 1-3, 2018; Atlanta, GA. [doi: 10.1109/ismr.2018.8333298]

129. Damodaran H, Adamovich S. Examining the manipulation of the dynamic properties of virtual objects to optimize upper extremity rehabilitation activities. 2010 Presented at: Proceedings of the 2010 IEEE 36th Annual Northeast Bioengineering Conference (NEBEC); March 26-28, 2010; New York, NY. [doi: 10.1109/nebc.2010.5458219]

130. Huber M, Rabin B, Docan C, Burdea GC, AbdelBaky M, Golomb MR. Feasibility of modified remotely monitored in-home gaming technology for improving hand function in adolescents with cerebral palsy. IEEE Trans Inf Technol Biomed 2010 Mar;14(2):526-534. [doi: 10.1109/TITB.2009.2038995] [Medline: 20071262]

131. Huber M, Rabin B, Docan C, Burdea G, Nwosu ME, Abdelbaky M, et al. PlayStation 3-based tele-rehabilitation for children with hemiplegia. 2008 Presented at: 2008 Virtual Rehabilitation; August 25-27, 2008; Vancouver, BC. [doi: 10.1109/icvr.2008.4625145]

132. Lioulemes A, Sassaman P, Gieser S, Karkaletsis V, Makedon F, Metsis V. Self-managed patient-game interaction using the barrett WAM arm for motion analysis. 2015 Presented at: PETRA '15: Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments; July 1-3, 2015; Corfu, Greece. [doi: 10.1145/2769493.2769517]

133. Cortés C, Ardanza A, Molina-Rueda F, Cuesta-Gómez A, Unzueta L, Epelde G, et al. Upper limb posture estimation in robotic and virtual reality-based rehabilitation. Biomed Res Int 2014;2014:821908. [doi: 10.1155/2014/821908] [Medline: 25110698]

134. Jee CHC, Perumal N, Elamvazuthi I, Tageldeen M, Khan M, Parasuraman S. Home-based interactive rehabilitation system for hand. 2016 Presented at: IEEE International Symposium on Robotics and Manufacturing Automation (ROMA). IEEE; September 25-27, 2016; Ipoh, Malaysia. [doi: 10.1109/roma.2016.7847835]

135. Padilla-Castañeda MA, Sotgiu E, Barsotti M, Frisoli A, Orsini P, Martiradonna A, et al. An Orthopaedic Robotic-Assisted Rehabilitation Method of the Forearm in Virtual Reality Physiotherapy. J Healthc Eng 2018;2018:7438609. [doi: 10.1155/2018/7438609] [Medline: 30154992]

136. Liu L, Chen X, Lu Z, Cao S, Wu D, Zhang X. Development of an EMG-ACC-Based Upper Limb Rehabilitation Training System. IEEE Trans Neural Syst Rehabil Eng 2017 Mar;25(3):244-253. [doi: 10.1109/tnsre.2016.2560906]

137. Ghassemi M, Triandafilou K, Barry A, Stoykov ME, Roth E, Mussa-Ivaldi FA, et al. Development of an EMG-Controlled Serious Game for Rehabilitation. IEEE Trans Neural Syst Rehabil Eng 2019 Feb;27(2):283-292 [FREE Full text] [doi: 10.1109/TNSRE.2019.2894102] [Medline: 30668478]

138. Wang Y, Su AWY, Han T, Lin C, Hsu L. EMG based rehabilitation systems - approaches for ALS patients in different stages. 2015 Presented at: 2015 IEEE International Conference on Multimedia and Expo (ICME); June 29-July 3, 2015; Turin, Italy. [doi: 10.1109/icme.2015.7177398]

139. Wu Q, Wang X, Chen B, Wu H. Patient-Active Control of a Powered Exoskeleton Targeting Upper Limb Rehabilitation Training. Front Neurol 2018;9:817. [doi: 10.3389/fneur.2018.00817] [Medline: 30364274]

140. Lipovský R, Ferreira HA. Self hand-rehabilitation system based on wearable technology. 2015 Presented at: REHAB '15: Proceedings of the 3rd 2015 Workshop on ICTs for improving Patients Rehabilitation Research Techniques; October 1-2, 2015; Lisbon, Portugal. [doi: 10.1145/2838944.2838967]

141. Batista TVV, Machado LS, Valenca AMG. Surface Electromyography for Game-Based Hand Motor Rehabilitation. 2016 Presented at: 2016 XVIII Symposium on Virtual and Augmented Reality (SVR); June 21-24, 2016; Gramado, Brazil. [doi: 10.1109/svr.2016.32]

142. Banerji S, Heng J, Pereira B. Upper extremity rehabilitation after stroke: biofeedback gaming for attention and muscle use. 2012 Presented at: i-CREATe '12: Proceedings of the 6th International Conference on Rehabilitation Engineering & Assistive Technology; July 24-26, 2012; Tampines, Singapore. [doi: 10.5555/2501134.2501172]

143. Lin BS, Hsu HC, Jan GE, Chen JL. An Interactive Upper-Limb Post-Stroke Rehabilitation System Integrating BCI-based Attention Monitoring and Virtual Reality Feedback. 2016 Presented at: 2016 Third International Conference on Computing Measurement Control and Sensor Network (CMCSN); May 20-22, 2016; Matsue, Japan. [doi: 10.1109/cmcsn.2016.33]

144. Lin B, Chen J, Hsu H. Novel Upper-Limb Rehabilitation System Based on Attention Technology for Post-Stroke Patients: A Preliminary Study. IEEE Access 2018;6:2720-2731. [doi: 10.1109/access.2017.2785122]

145. Castellini C, Hertkorn K, Sagardia M, Gonzalez DS, Nowak M. A virtual piano-playing environment for rehabilitation based upon ultrasound imaging. 2014 Presented at: 5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics; August 12-15, 2014; Sao Paulo, Brazil. [doi: 10.1109/biorob.2014.6913835]

146. Jie S, Haoyong Y, Chaw TL, Chiang CC, Vijayavenkataraman S. An Interactive Upper Limb Rehab Device for Elderly Stroke Patients. Procedia CIRP 2017;60:488-493. [doi: 10.1016/j.procir.2017.02.040]

147. Rahman HA, Kamal AAA, Narayanan ALT. A simple upper limb rehabilitation trainer. Int J Integr Eng 2017;9(3):39-43.

148. Dunne A, Do-Lenh S, Laighin G, Shen C, Bonato P. Upper extremity rehabilitation of children with cerebral palsy using accelerometer feedback on a multitouch display. 2010 Presented at: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology; August 31-September 4, 2010; Buenos Aires, Argentina. [doi: 10.1109/iembs.2010.5626724]

149. Wang H, Hsu C, Chiu D, Tsai SN. Using augmented reality gaming system to enhance hand rehabilitation. 2010 Presented at: 2010 2nd International Conference on Education Technology and Computer; June 22-24, 2010; Shanghai, China. [doi: 10.1109/icetc.2010.5529553]

150. Liu S, Meng D, Cheng L, Huang F. A Virtual Reality based Training and Assessment System for Hand Rehabilitation. 2018 Presented at: 2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP); November 9-11, 2018; Wanzhou, China. [doi: 10.1109/icicip.2018.8606705]

151. Basteris A, Johansson E, Klein P, Nasr N, Nijenhuis S, Sale P, et al. Creating gesture controlled games for robot-assisted stroke rehabilitation. Biomed Tech 2014;59(Suppl 1):S706-S709. [doi: 10.1515/bmt-2014-4301]

152. Amirabdollahian F, Ates S, Basteris A, Cesario A, Buurke J, Hermens H, et al. Design, development and deployment of a hand/wrist exoskeleton for home-based rehabilitation after stroke - SCRIPT project. Robotica 2014 Sep 23;32(8):1331-1346. [doi: 10.1017/S0263574714002288]

153. Shah N, Amirabdollahian F, Basteris A. Designing motivational games for stroke rehabilitation. 2014 Presented at: 2014 7th International Conference on Human System Interactions (HSI); June 16-18, 2014; Costa da Caparica, Portugal. [doi: 10.1109/hsi.2014.6860468]

154. Rosman R, Abdul Hadi MZ, Abu Bakar N. Interactive Design and Development of Real Arm Movements for Application in Rehabilitation. 2018 Apr 09 Presented at: IOP Conference Series: Materials Science and Engineering; August 8, 2017; Kuching, Sarawak, Malaysia p. 012007. [doi: 10.1088/1757-899x/341/1/012007]

155. Friedman N, Chan V, Zondervan D, Bachman M, Reinkensmeyer DJ. MusicGlove: Motivating and quantifying hand movement rehabilitation by using functional grips to play music. 2011 Presented at: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 30-September 3, 2011; Boston, MA. [doi: 10.1109/iembs.2011.6090659]

156. Anam K, Rosyadi AA, Sujanarko B. The Design of a Low-Cost Therapy Robot for Hand Rehabilitation of a Post-Stroke Patient. 2018 Presented at: 2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM); November 26-27, 2018; Surabaya, Indonesia. [doi: 10.1109/cenim.2018.8710833]

157. Zheng J, Shi P, Yu H. A Virtual Reality Rehabilitation Training System Based on Upper Limb Exoskeleton Robot. 2018 Presented at: 2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC); August 25-26, 2018; Hangzhou, China. [doi: 10.1109/ihmsc.2018.00058]

XSL•FO
RenderX

158. Rodriguez-de-Pablo C, Perry JC, Cavallaro FI, Zabaleta H, Keller T. Development of computer games for assessment and training in post-stroke arm telerehabilitation. 2012 Presented at: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 28-September 1, 2012; San Diego, CA. [doi: 10.1109/embc.2012.6346984]

159. Rodriguez-de-Pablo C, Perry JC, Balasubramanian S, Belloso A, Savic A, Tomic TD, et al. Serious Games for Assessment and Training in Post-stroke Robotic Upper-limb Telerehabilitation. 2014 Presented at: Proceedings of the 2nd International Congress on Neurotechnology, Electronics and Informatics; October 25-26, 2014; Rome, Italy. [doi: 10.5220/0005168601260134]

160. Simkins M, Fedulow I, Kim H, Abrams G, Byl N, Rosen J. Robotic Rehabilitation Game Design for Chronic Stroke. Games Health J 2012 Dec;1(6):422-430. [doi: 10.1089/g4h.2012.0044] [Medline: 26192059]

161. Tutak J. Virtual reality and exercises for paretic upper limb of stroke survivors. Teh Vjesn 2017 Sep;24(Suppl 2):451-458. [doi: 10.17559/tv-20161011143721]

162. Agas AM, Daitol AA, Shah UB, Fraser LJ, Abbruzzese KM, Karunakaran KK, et al. 3-DOF Admittance Control Robotic Arm with a 3D Virtual Game for facilitated training of the hemiparetic hand. 2015 Presented at: 2015 41st Annual Northeast Biomedical Engineering Conference (NEBEC); April 17-19, 2015; Troy, NY. [doi: 10.1109/nebec.2015.7117071]

163. Padilla-Castaneda MA, Sotgiu E, Frisoli A, Bergamasco M, Orsini P, Martiradonna A, et al. A virtual reality system for robotic-assisted orthopedic rehabilitation of forearm and elbow fractures. 2013 Presented at: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems; November 3-7, 2013; Tokyo, Japan. [doi: 10.1109/iros.2013.6696548]

164. Li S, Frisoli A, Avizzano C, Ruffaldi E, Lugo-Villeda L, Bergamasco M. Bimanual Haptic-desktop platform for upper-limb post-stroke rehabilitation: Practical trials. 2009 Presented at: 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO); December 19-23, 2009; Guilin, China. [doi: 10.1109/robio.2009.5420694]

165. Weightman APH, Preston N, Holt R, Allsop M, Levesley M, Bhakta B. Engaging children in healthcare technology design: developing rehabilitation technology for children with cerebral palsy. J Eng Design 2009 Mar 25;21(5):579-600. [doi: 10.1080/09544820802441092]

166. Sivak M, Unluhisarcikli O, Weinberg B, Mirelman-Harari A, Bonato P, Mavroidis C. Haptic system for hand rehabilitation integrating an interactive game with an advanced robotic device. 2010 Presented at: 2010 IEEE Haptics Symposium; March 25-26, 2010; Waltham, MA. [doi: 10.1109/haptic.2010.5444615]

167. Tanaka Y, Matsushita K, Tsuji T, Imamura N. Human Hand Movements in Skilled Motor Performance of a Virtual Curling Task. 2007 Presented at: 2007 IEEE/ICME International Conference on Complex Medical Engineering; May 23-27, 2007; Beijing, China. [doi: 10.1109/iccme.2007.4381957]

168. Peng L, Hou ZG, Kasabov N, Peng L, Hu J, Wang W. Implementation of active training for an upper-limb rehabilitation robot based on impedance control. 2015 Presented at: The 27th Chinese Control and Decision Conference (2015 CCDC); May 23-25, 2015; Qingdao, China. [doi: 10.1109/ccdc.2015.7161769]

169. Yamabe K, Sato C, Abe I, Kikuchi T, Inoue A, Kugimiya S, et al. Motor function assessment for upper limb with virtual wiping program of rehabilitation system SEMUL. 2015 Presented at: 2015 IEEE/SICE International Symposium on System Integration (SII); December 11-13, 2015; Nagoya, Japan. [doi: 10.1109/sii.2015.7404891]

170. Jiang N, Dai Y, Li X, Jin Y. Research on virtual reality based on RRR-I. 2012 Presented at: 2012 24th Chinese Control and Decision Conference (CCDC); May 23-25, 2012; Taiyuan, China. [doi: 10.1109/ccdc.2012.6244457]

171. Azzam K, Takrouri M, Abu-Alfielat F, Ideis M, Tahboub K. Virtual rehabilitation robot (VRR). 2013 Presented at: 2013 9th International Symposium on Mechatronics and its Applications (ISMA); April 9-11, 2013; Amman, Jordan. [doi: 10.1109/isma.2013.6547394]

172. Shao W, Yu H, Xie L. A convenient home-based rehabilitation system for patients with paretic upper-limb. 2011 Presented at: i-CREATe '11: Proceedings of the 5th International Conference on Rehabilitation Engineering & Assistive Technology; July 2011; Bangkok, Thailand p. 243-245. [doi: 10.5555/2500753.2500811]

173. Chang C, Chang Y, Chang H, Chou L. An interactive game-based shoulder wheel system for rehabilitation. Patient Prefer Adherence 2012;6:821-828. [doi: 10.2147/PPA.S37190] [Medline: 23226005]

174. Cifuentes-Zapien J, Valdez-Aguilar J, Rojas-Correa F, Chong-Quero J, Pineda-Olivares A. A video game for an upper limb rehabilitation robotic system for children with cerebral palsy. 2011 Presented at: 2011 Pan American Health Care Exchanges; March 28-April 1, 2011; Rio de Janeiro, Brazil. [doi: 10.1109/pahce.2011.5871877]

175. Takahashi Y, Terada T, Inoue K, Ito Y, Ikeda Y, Lee H, et al. Haptic Device System for Upper Limb Motor and Cognitive Function Rehabilitation: Grip Movement Comparison between Normal Subjects and Stroke Patients. 2007 Presented at: 2007 IEEE 10th International Conference on Rehabilitation Robotics; June 13-15, 2007; Noordwijk, Netherlands. [doi: 10.1109/icorr.2007.4428507]

176. Jordan K, Sampson M, Hijmans J, King M, Hale L. ImAble system for upper limb stroke rehabilitation. 2011 Presented at: 2011 International Conference on Virtual Rehabilitation; June 27-29, 2011; Zurich, Switzerland. [doi: 10.1109/icvr.2011.5971835]

177. Kowalczewski J, Ravid E, Prochazka A. Fully-automated test of upper-extremity function. 2011 Presented at: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 30-September 3, 2011; Boston, MA. [doi: 10.1109/iembs.2011.6091710]

178. Levin MF. Interjoint coordination during pointing movements is disrupted in spastic hemiparesis. Brain 1996 Feb;119(Pt 1):281-293. [doi: 10.1093/brain/119.1.281] [Medline: 8624689]

179. Osu R, Ota K, Fujiwara T, Otaka Y, Kawato M, Liu M. Quantifying the quality of hand movement in stroke patients through three-dimensional curvature. J Neuroeng Rehabil 2011 Oct 31;8:62 [FREE Full text] [doi: 10.1186/1743-0003-8-62] [Medline: 22040326]

180. World Health Statistics 2012. World Health Organization. 2012. URL: https://www.who.int/gho/publications/world_health_statistics/EN_WHS2012_Full.pdf [accessed 2019-03-01]

181. Coates G. Program from Invisible Site-a Virtual Sho. In: a Multiple Performance Work. San Francisco: George Coates Performance Works; Mar 1992.

Abbreviations

AR: augmented reality
ARAT: Action Research Arm Test
EMG: electromyogram
FMA: Fugl-Meyer Motor Function Assessment
IMU: inertial measurement unit
MAUULF: Melbourne Assessment of Unilateral Upper Limb Function
mHealth: mobile health
ROM: range of motion
VR: virtual reality

Original Paper

Effect of Spatial Disorientation in a Virtual Environment on Gait and Vital Features in Patients with Dementia: Pilot Single-Blind Randomized Control Trial

Chimezie O Amaefule[1], MSc; Stefan Lüdtke[2], MSc; Thomas Kirste[2], DEng; Stefan J Teipel[1,3], DMed

[1]German Center for Neurodegenerative Diseases (DZNE), Rostock, Germany

[2]Institute of Visual & Analytic Computing, University of Rostock, Rostock, Germany

[3]Department of Psychosomatic and Psychotherapeutic Medicine, University Medicine Rostock, Rostock, Germany

Corresponding Author:
Chimezie O Amaefule, MSc
German Center for Neurodegenerative Diseases (DZNE)
Gehlsheimer Strasse 20
Rostock, 18147
Germany
Phone: 49 381 494 9478
Fax: 49 381 494 9472
Email: chimezie.amaefule@dzne.de

Abstract

Background: Orientation deficits are among the most devastating consequences of early dementia. Digital navigation devices could overcome these deficits if adaptable to the user's needs (ie, provide situation-aware, proactive navigation assistance). To fulfill this task, systems need to automatically detect spatial disorientation from sensors in real time. Ideally, this would require field studies consisting of real-world navigation. However, such field studies can be challenging and are not guaranteed to cover sufficient instances of disorientation due to the large variability of real-world settings and a lack of control over the environment.

Objective: Extending a foregoing field study, we aim to evaluate the feasibility of using a sophisticated virtual reality (VR) setup, which allows a more controlled observation of disorientation states and accompanying behavioral and physiological parameters in cognitively healthy older people and people with dementia.

Methods: In this feasibility study, we described the experimental design and pilot outcomes of an ongoing study aimed at investigating the effect of disorientation on gait and selected physiological features in a virtual laboratory. We transferred a real-world navigation task to a treadmill-based virtual system for gait analysis. Disorientation was induced by deliberately manipulating landmarks in the VR projection. Associated responses in motion behavior and physiological parameters were recorded by sensors. Primary outcomes were variations in motion and physiological parameters, frequency of disorientation, and questionnaire-derived usability estimates (immersion and perceived control of the gait system) for our population of interest. At this time, the included participants were 9 cognitively healthy older participants [5/9 women, 4/9 men; mean age 70 years, SD 4.40; Mini–Mental State Examination (MMSE) mean 29, SD 0.70) and 4 participants with dementia (2/4 women, 2/4 men; mean age 78 years, SD 2.30 years; MMSE mean 20.50, SD 7.54). Recruitment is ongoing, with the aim of including 30 cognitively healthy older participants and 20 participants with dementia.

Results: All 13 participants completed the experiment. Patients' route was adapted by shortening it relative to the original route. Average instances of disorientation were 21.40, 36.50, and 37.50 for the cognitively healthy older control, cognitively healthy older experimental participants, and participants with dementia, respectively. Questionnaire outcomes indicated that participants experienced adequate usability and immersion; 4.30 for presence, 3.73 for involvement, and 3.85 for realism of 7 possible points, indicating a good overall ability to cope with the experiment. Variations were also observed in motion and physiological parameters during instances of disorientation.

Conclusions: This study presents the first feasibility outcomes of a study investigating the viability of using a sophisticated VR setup, based on an earlier real-world navigation study, to study spatial disorientation among cognitively healthy older people and people with dementia. Preliminary outcomes give confidence to the notion that our setup can be used to assess motion and physiological markers of disorientation, even in people with cognitive decline.

Trial Registration: ClinicalTrials.gov; https://clinicaltrials.gov/ct2/show/NCT04134806

XSL•FO
RenderX

Introduction

Background

Challenges in wayfinding and orientation are early symptoms of people with mild cognitive impairment (MCI) or dementia. These deficits decrease mobility [1,2] and social interaction [3,4] of the affected people, which in turn may lead to further cognitive decline. Assistive technology devices (ATDs) can help reduce the burden of spatial disorientation by providing interventions (eg, by giving cues to the patients). For patients with cognitive impairment, an ideal ATD should fulfill 2 requirements: It should be situation-adaptive (ie, the device adapts to the situation and context, including the environment, user goal, and intention), and it should be subsidiary (ie, the device delivers assistance only in case of need, for example, if the user is disoriented). These requirements ensure that the device does not replace but rather leverages existing cognitive capabilities [5]. Technically, this means that the ATD needs to be able to detect instances of disorientation in real time from the available sensor data, like accelerometric, electrodermal (EDA), or electrocardiographic (ECG) data.

In a previous field study—Situation-Aware Navigation Assistance for Dementia Patients using Causal Behavior Models (SiNDeM) [6]—concerned with wayfinding behavior in MCI and patients with dementia through an urban environment, a machine learning classifier for disorientation based solely on accelerometric data led to a cross-validated area under the receiver operating characteristics curve (AUC) value of 0.75. The outcome of this study suggested that instantaneous detection of disorientation may, in principle, be possible; however, the accuracy was not sufficient to serve as a basis for individual support. Hence, additional signals, such as heart rate variability, may be needed to increase detection accuracy. However, performing such a study in a real-world environment requires a large effort in staff and resources. One of the limitations highlighted in the previous study [6] was that the recorded number of disorientation instances per subject was low, which can be problematic when training machine learning classifiers. This is due to the fact that, firstly, the experimenter could not influence whether (and when) subjects became disoriented. Secondly, the high level of inconsistency in the real-world environment did not enable the controlled observation of disorientation states that are not induced by experimental manipulation (eg, changing of landmarks) but rather by cognitive deficits, as in the case of the patients. Thus, a large effort is required to obtain only a small amount of data related to disorientation instances (which is the data that is most relevant in this context). A more robust approach to modeling real-time disorientation might rely on both a controlled environment as well as other more informative parameters in addition to motion. Gait features have recently been vastly explored as motion markers of disease progression [7-9] and fall detection [10], and could also be informative in identifying behavioral variations predictive of disorientation. Additionally, physiological parameters, including skin conductance response and heart rate variations, have also been previously shown to be influenced by the occurrence of spatial disorientation [11,12].

As an alternative to real-world studies, navigation tasks can be posed in a virtual-reality (VR) environment [13,14]. However, most experimental setups do not integrate a physical component—participants sit in front of a computer screen, and thus the physical manifestation of disorientation cannot be assessed. More generally, the need for physical locomotion might influence navigation behavior, as participants are in a dual-task situation where they have to simultaneously walk and navigate through the environment. Such dual-task conditions require cognitive resources such as attentional flexibility, which are depleted as disease progresses [15-18]. Therefore, a pertinent question would be, how can the navigation task be transferred to a safe and controllable VR environment while the participants still have to walk actively (like they would do in the real world), to allow for the investigation of the relationship between disorientation and physical motion? To this extent, we propose a VR-based experimental setup that allows realistic physical movement.

Specifically, we employed the GRAIL (Gait Real-time Analysis Interactive Lab; Motekforce Link) system, providing the opportunity to navigate on a treadmill through a virtual environment. Using such a virtual environment has several advantages: The setup is safe for the subjects, reproducible, and the experimental effort per subject is lower compared to a real-world study. Furthermore, the experimenters have full control over the environment; for example, the environment can be manipulated to induce disorientation to record a larger amount of disorientation instances. Also, disorientation states resulting from cognitive deficits can be properly observed.

Objectives

This study aims to evaluate the feasibility of using a complex virtual reality setup, which allows a more controlled observation of disorientation states, to study spatial disorientation among older cognitively healthy people and people with dementia, thereby validating the findings of the real-world SiNDeM study. This environment gives access to a broader set of sensor domains compared with the real-world setting, including gait, skin conductance, and heart rate variability in addition to accelerometry; it also allows for the examination of a larger sample size within a more controlled environment to build a more accurate disorientation detection and intervention model.

Research Questions

The study is based on the following core research question: Is our setup feasible for investigating states of disorientation during active navigation among cognitively healthy older participants and people with dementia? This is motivated by the reported overall effectiveness of VR in assessing spatial navigation [19],

spatial navigation memory for predementia screening [20], and improving cognitive functioning among individuals with neurocognitive disorders [21,22]. Evaluation of the feasibility of our setup is further guided by the following research questions: (1) Is the virtual environment adequately immersive? (2) Are we able to reliably induce disorientation through the manipulation of the virtual environment? (3) Do the participants feel comfortable with the walking pattern change due to our navigation mechanism? (4) Are we able to adequately measure motion and physiological parameters?

Methods

Study Design and Setting

This is a single-blind randomized experiment currently taking place in Rostock, Germany. The experiments are carried out in the Gait Real-time Analysis Interactive Lab (GRAIL; Motekforce Link; Figure 1), which is a Class I medical system (according to the Medical Device Directive 93/42/EEC) and is specially designed to ensure the safety of participants for clinical gait analysis and training. This is achieved through a safety belt and side railings, which serve to support the participant and prevent falls. The GRAIL further consists of a treadmill, a large 180° projection screen, and an optical motion capturing system (Vicon Motion Systems Ltd). A virtual environment that models a city center is shown onscreen, and subjects can navigate through the environment by walking on the treadmill. Gait kinematics and kinetics, as well as spatio-temporal gait parameters, can be derived from the motion capturing data [23]. The system has been used, for example, for rehabilitation exercise [24] or for gait analysis in different settings [25-27].

Figure 1. (A) The GRAIL system, consisting of a dual-belt treadmill, 180° projection surface, and optical motion capturing system (figure sourced from Motek Medical); (B) plantar figures.
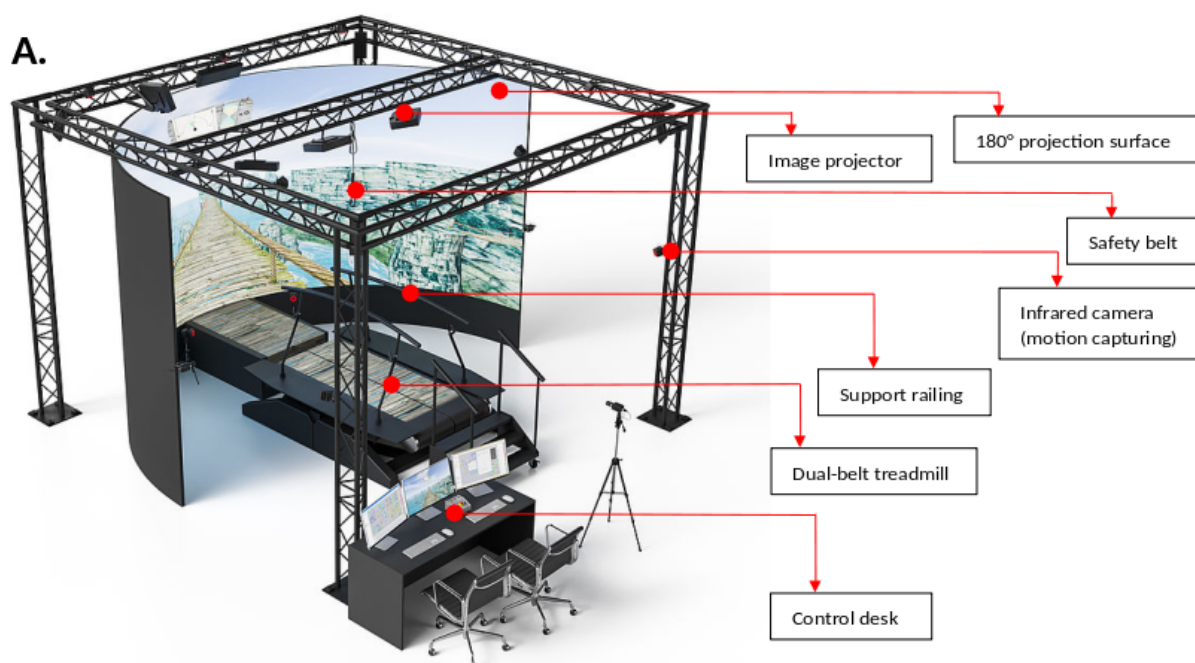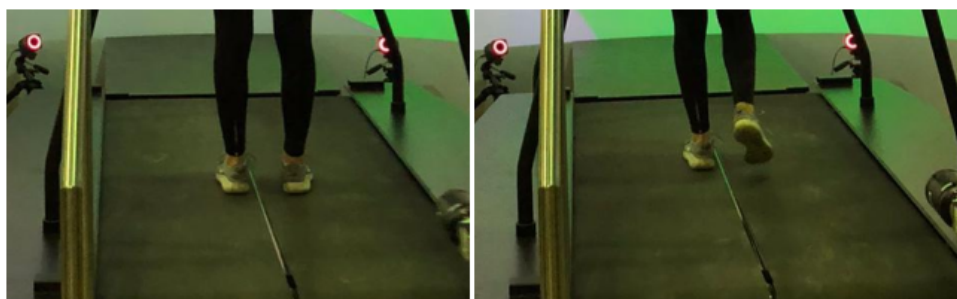
The 3D virtual environment used in this study was generated from OpenStreetMap (OSM) data of the Rostock city center, using the OSM2World tool [28]. This data includes building heights and rudimentary 3D models of landmark buildings (Figure 2). The resulting VR environment is a low-detail replication of the real city, but does not contain moving objects like cars or pedestrians. The treadmill speed is feedback-controlled, which allows participants to walk with their preferred walking speed. This is achieved by adapting the belt speed according to the subject position (measured by motion capturing) in relationship to the center of the belt [29]. The movement speed through the VR is synchronized with the belt

speed (ie, the current speed of the participant). An important constraint of the GRAIL system is that a change in walking direction (eg, turning left or right) is not supported by the treadmill; it does not rotate. Thus, it was necessary to provide an alternative means for voluntary direction change: participants can choose their walking direction by walking on either side of the treadmill. Walking on the left side of the treadmill will result in a left turn in the VR environment, and vice versa. We are aware that this is not identical to naturalistic walking: subjects still walk a straight line on the treadmill, whereas the movement in the VR describes a curve. Therefore, as part of the current study reports, we observed the movement behavior of participants and asked them about their experiences.

Figure 2. (A) The OpenStreetMap (OSM)–generated virtual environment containing some notable landmarks from (B) the real world.
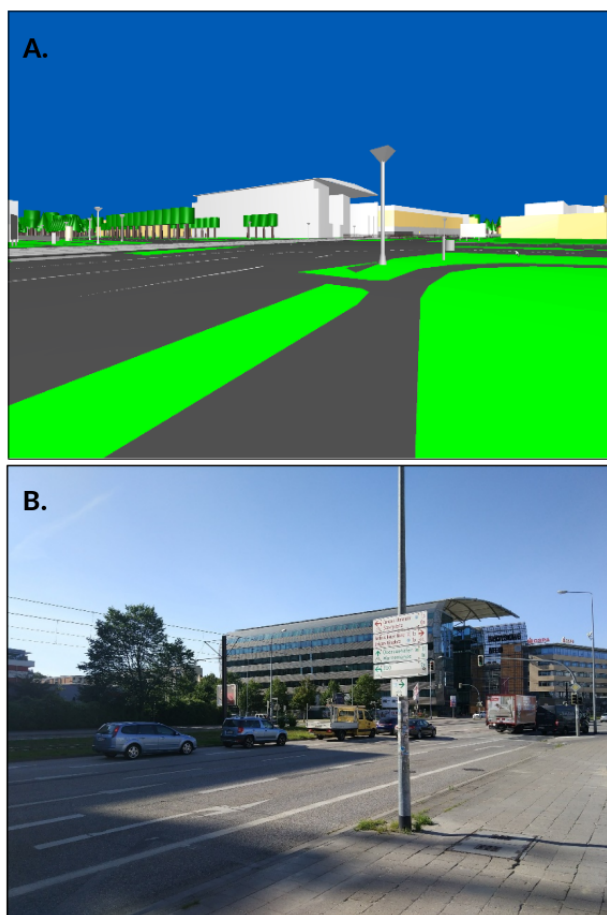
Recruitment and Eligibility

Participants were recruited in 3 groups: (1) mobile, physically and cognitively healthy, younger (18-40 years of age) adults; (2) mobile, physically healthy older [60-85 years of age; Mini–Mental State Examination (MMSE) scores≥28] adults, including cognitively healthy older adults without memory complaints and cognitively healthy older adults with subjective cognitive decline (SCD) in the absence of any clinical evidence of cognitive impairment; and (3) physically healthy older adults with diagnosed MCI or mild dementia due to Alzheimer disease (60-85 years of age; MMSE 15-27). People with dementia and cognitively healthy older adults are recruited from the memory clinic of the University Medicine, Rostock, while the healthy young adults are recruited from within the University of Rostock student community. Exclusion criteria for all groups include other neurological conditions besides dementia, an inability to understand task instructions and questionnaire items, and deaf-mutism and blindness.

As the focus of this feasibility study was on cognitively healthy older adults and people with dementia, 9 cognitively healthy older participants (5/9 women, 4/9 men; 6/9 with SCD; mean age 70 years, SD 4.40; MMSE mean 29, SD 0.70) and 4 people with dementia (2/4 women, 2/4 men; mean age 78 years, SD 2.30; MMSE mean 20.50, SD 7.54) have been included so far. Of the 9 cognitively healthy older participants, 4 participants (2/4 women, 2/4 men) were randomly assigned to the experimental group. Informed consent was given by all participants.

Study Procedure and Data Collection

Participants were guided along a path in the VR environment (Figure 3). Afterward, they were set back to the starting location and asked to walk the same path again, this time unguided. For half of the healthy young or older subjects (the experimental group), phases of disorientation were induced by changing landmarks or decision points in the VR environment (Figure 4) while subjects were required to walk the path on their own. The changes were (1) moving a landmark from one intersection to the next intersection, (2) adding a decision point (ie, an

intersection), (3) blocking a road, or (4) moving the goal indicator to a different location. Overall, 5 locations have been manipulated. For the patients with dementia, the environment was not changed, as we assume that these participants would show phases of disorientation already without such changes due

to existing cognitive deficits. Based on the experience with the first patient with dementia, we adapted the route for the other patients by shortening it due to earlier observed fatigue while navigating the route (Figure 3).

Figure 3. Map of the routes of (A) cognitively healthy older participants and (B) patients with dementia. Red crosses denote locations where the environment is changed in the experimental run. The environment is unchanged for the patients with dementia; however, the route is shortened (image: Google Maps).

Figure 4. Example of changes in the environment to induce disorientation. (A) Original environment shown in the guided walk (note the red landmark at the far end of the road, to the right). (B) Manipulated environment used in the experimental run; in this case, the landmark is moved to a different intersection.
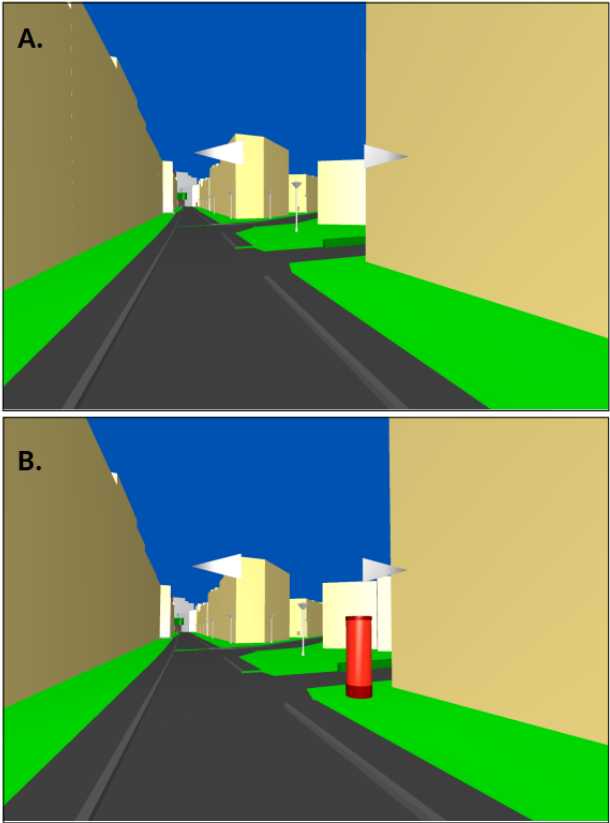
The participants were familiarized with the depicted city center by briefly showing them a map, such that problems in wayfinding will be due to disorientation instead of exploration in an unknown environment. In addition to recording kinematic and kinetic gait parameters as provided by the GRAIL system, participants are also equipped with 3 wearable sensors on the left wrist, right ankle, and chest, each of which contain a 3-axes accelerometer and 3-axes gyroscope sampled with 64 Hz. Additionally, the chest sensor records an electrocardiogram (ECG; 1024 Hz), and the wrist sensor records electrodermal activity (EDA; 32 Hz). Wearable sensor data and data provided by the GRAIL system are synchronized by an event-based mechanism (ie, participants perform a distinctive movement at the beginning of the recording, which can be easily located in all sensors) and are resampled to 100 Hz. All wearable sensors have been used in previous studies [6,30] and validated [31-33].

After the orientation task, participants filled out the igroup Presence Questionnaire (IPQ) [34], a questionnaire on the functionality and realism of the interaction with the VR environment, as well as questions regarding engagement in the study, hobbies and activities, technical device usage, and demographics. Additionally, participants provided answers to open questions on (1) problems that occurred in controlling the direction, (2) suggestions for improving the controlling mechanism, (3) which properties in the VR they used for orientation, (4) orientation problems they had, and (5) suggestions for enabling better self-orientation in the VR environment. Table 1 displays the experimental procedure in phases.

Table 1. Experimental procedure for the cognitively healthy older participants (n=9) and patients with dementia (n=4). The patient group followed the procedure of the control group.

	Phase	Experimental group (n=4)	Patients with dementia (n=4) and control group (n=5)
1	Preparation (90 min)	Study information, informed consent, assessment of physical and cognitive status, blood draw, fixing of markers and electrodes, and practice walking on treadmill	
2	Task 1 (20 min)	Learning the route: accompanied walk (route: Figure 3)	
3	Task 2 (20 min)	Autonomous navigation (modified environment)	Autonomous navigation (unchanged environment)
4	Questionnaires (20 min)	Presence, navigation, orientation, experience with technical devices	

Randomization

Randomization of the healthy younger and older participants into the experimental or control group was carried out using the program Research Randomizer (Social Psychology Network) [35]. In contrast, the patients with dementia are only assigned to the control group (ie, with the adapted route), as we expect a sufficient number of episodes of disorientation in people with MCI or dementia even without interfering with their environment.

Behavior Annotation

An offline annotation procedure was applied to the video data recorded during the orientation task, for assessing the observable orientation behavior of the participants using the ELAN 5.8 tool (The Language Archive) [36]. As a coding scheme, we used an adequate adaption of the coding scheme provided by Yordanova et al [30]. The same scheme has been used in the SiNDeM field study [6]. This coding scheme was developed both by domain experts and assistive systems designers, based on interviews, video logs, data from a systematic literature review, and concepts from existing ontologies, for the purpose of providing assistance to people with dementia during their outdoor mobility. Hence, it covers aspects of orientation behavior that are beyond the scope of wayfinding in our VR setup (eg, behaviors associated with attention to traffic). For this reason, we adapted the coding scheme to capture exactly those behaviors that are obtainable within our virtual reality setup.

Specifically, to identify instances of disorientation, we annotated when participants show wandering behavior (ie, nongoal-directed walk), communication behavior (ie, asking for help when disoriented), topological orientation (ie, trying to orient themselves based on the surrounding environment), or spatial orientation (ie, trying to orient themselves based on landmarks). In addition, different types of errors that are associated with disoriented behavior were annotated (ie, initiation, realization, sequence, and completion errors). The annotations are being evaluated based on the level of agreement between annotators (ie, interrater reliability in terms of Cohen kappa [37]).

Ethical Approval

This study has been reviewed and approved by the Ethics Commission of the University Medicine Rostock (Approval number: A 2019-0062).

Outcome Measures

The study is estimated to run until the end of 2021. The outcome measures to be collected are included in Table 2. However, for this pilot study, we focused on the feasibility and usability (level of immersion, perceived control of the gait system) as primary outcomes, and on measures of motion variations (walking speed), physiological variations (heart rate, skin conductance response), and spatial disorientation (frequency of occurrence) as secondary outcomes.

Table 2. Study outcome measures.

Outcome Measure	Measurement	Modality	Status
Feasibility and usability	Level of immersion, usability feedback	igroup Presence Questionnaire (IPQ), usability questionnaire	Ongoing
Heart rate variability	Rate of change in heart rate	Electrocardiographic sensor	Ongoing
Skin conductance	Rate of change in electrodermal response	Electrodermal activity sensor	Ongoing
Spatial disorientation	Incidences of disorientation	Customized annotation scheme in ELAN	Ongoing
Gait variability	Incidences of change in gait pattern	Gait capturing system of the GRAIL	Ongoing
Accelerometry	Incidences of change in motion pattern	Accelerometers	Ongoing
Apolipoprotein E4 status	Presence of the variants Apo-E2, -E3, and -E4 in the blood samples	7.5 ml blood samples	Ongoing

Data Analysis

Analysis of both the quantitative and qualitative data collected during the course of the study will take place in accordance with predetermined analysis plans; however, for this feasibility study with a limited sample size, we reported basic descriptive statistics (mean and standard deviation as well as frequencies) applied to the quantitative and qualitative data for evaluation purposes using R statistical software (version 3.6.0; R Core Team).

Results

Immersiveness of the Virtual Environment

All participants could complete the experiment. Responses to the IPQ informed us that the cognitively healthy older participants perceived an above-average degree of immersion; group mean item scores (between 1 and 7, where 7 means highest perceived presence/involvement/realism) were 4.60 for presence, 3.92 for involvement, and 4.42 for realism. The patients with dementia, on the other hand, while reporting lower mean scores relative to the cognitively healthy older participants, still perceived a considerable degree of immersion, with group mean item scores of 3.63 for presence, 3.31 for involvement,

and 2.60 for realism. Of the 13 participants, only 1 (7%) of the participants reported simulator sickness.

Inducing Disorientation Through the Manipulation of the Virtual Environment

Our setup was viable in inducing instances of disorientation. We observed an average of 21.40 instances of disorientation for the cognitively healthy older participants in the control group and 36.50 instances for the cognitively healthy older participants in the experimental group. For the patients with dementia, an average of 37.50 instances of disorientation was observed. A number of these instances of disorientation were observed either at points where the virtual environment was manipulated or at subsequent points afterward, where the participants had to reorient themselves due to the altered virtual environment (Figure 4). This amounted to a good proportion of the data being annotated as disoriented. Furthermore, regarding properties used for orientation, participants mentioned the landmark shown in Figure 4, houses, intersections, and trees.

Comfortability With the Walking Pattern Change due to our Navigation Mechanism

Participants' responses to the usability questionnaire show that the control over the chosen direction in the VR environment was perceived as functional for both the cognitively healthy older participants and patients with dementia (eg, easy to learn; participants were able to move to where they wanted at their own pace), and adequately naturalistic. Table 3 shows the answer scores on the questionnaire items regarding usability and navigation control. When asked about problems that occurred, only 2 (22%) of the 9 cognitively healthy older participants mentioned initial difficulty with controlling the direction of movement, while 1 (25%) of the 4 patients with dementia mentioned that it felt unusual. Of the 9 cognitively healthy older participants, 2 (22%) also mentioned that the right and left movement felt a little rapid. Additionally, in the case of the first patient with dementia sampled, we observed early fatigue while navigating the route, leading to an adaptation of the patient's route (Figure 3). Also, due to a temporary technical fault at the time of recording, we could not obtain gait information for the first patient with dementia sampled. However, all other data were collected for this patient, and complete data could be collected for the 3 subsequent patients with dementia sampled.

Table 3. Questionnaire regarding the usability of controlling movement in the virtual-reality environment for the cognitively healthy older participants (n=9; group mean score) and patients with dementia (n=4; group mean score). The scale ranged from 1 (fully disagree) to 7 (fully agree).

Questionnaire item	Cognitively healthy older participants (n=9), mean score	Patients with dementia (n=4), mean score
1. Learning to navigate was easy.	4.9	4.0
2. After a while, I did not have to think about how to navigate.	4.3	3.0
3. I was able to move where I wanted.	5.8	3.6
4. I was able to move how I wanted	4.9	3.0
5. I was able to stop at a specific place when I wanted.	5.9	4.0
6. I did not feel limited in my freedom of movement.	5.8	3.3
7. I felt tired after navigating.	3.3	2.5
8. Navigation felt natural.	4.7	3.3

Measuring Motion and Physiological Parameters

Variations in walking speed, heart rate, and skin conductance response were observed among our pilot participants. Figures 5A, B, and C, respectively, show a sample of the recorded walking speed, accelerometry, heart rate variability, and skin conductance response amplitude for a participant in the control group (where the environment was not manipulated), a participant in the experimental group (where the environment was manipulated, as described above), and for a patient with dementia (where the environment was not manipulated). Relative decrease in motion (walking speed and accelerometry) and an increase in heart rate variability and skin conductance response amplitude can be observed during instances of disorientation.

Figure 5. Sample of motion and physiological data showing variations in walking speed, accelerometry, heart rate, and skin conductance of participants occurring during instances of disorientation (green bars in the first row) from the unguided walk for (A) an older control participant (environment not manipulated), (B) an older experimental participant (environment was manipulated), and (C) a patient with dementia (environment not manipulated).
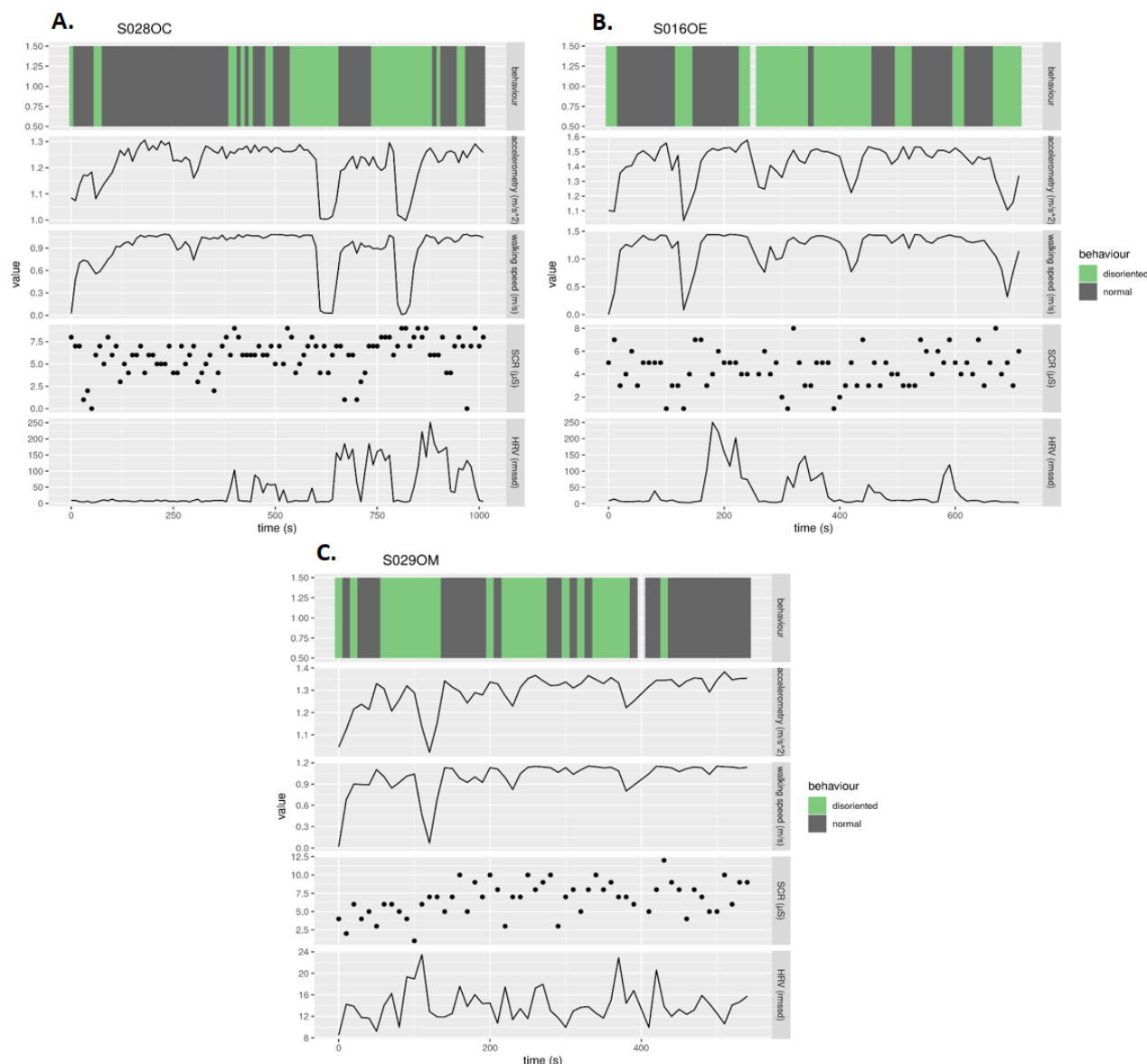
Discussion

In this paper, we presented the design for a study investigating the effect of spatial disorientation on physical motion and physiological features, as well as the initial outcomes of the pilot sample. The combined usage of a fully instrumented treadmill, wearable sensors, and an adequately immersive VR system to investigate real-time disorientation among older participants and patients with dementia during active navigation is, to the best of our knowledge, unprecedented. Previous studies that investigated spatial orientation [38-40] and navigation memory [20,41] have relied mostly on VR systems in which navigation or interaction with the VR environment was passively based on the use of control objects (eg, joystick, mouse). This, however, further limits the ecological validity of such setups, as the locomotion aspect of real-world navigation was apparently lacking.

Our patient and cognitively healthy older participant dyad reported a noticeably varying degree of immersion. While the older healthy participant group reported a higher level of immersion [42], the patient group reported a relatively lower level of immersion compared to the cognitively healthy older participant group. However, this is not surprising as we also expect cognitive and perceptual abilities to play a significant role in the evaluation of the VR setup, as people with dementia are known to experience visuoperceptual difficulties [43,44], which may influence this judgment. Nonetheless, participants' responses to the usability part of the questionnaire indicated an overall confidence that our setup is effective in investigating manifestations of disorientation among cognitively healthy older adults and people with dementia. Furthermore, the first patient with dementia sampled experienced fatigue while navigating the route (akin to the finding of Behrens et al [27]), which suggests that dual-task conditions can be mentally exhausting among older participants. Therefore, this effect could be

exacerbated by the increased difficulty of multitasking as a result of cognitive deficits [45]. Hence, we adapted the route of the patients by shortening it relative to the original route.

Considering the number of instances of spatial disorientation, we confirmed our expectation of finding more disorientation events among the participants in the experimental group (changed environment) compared to the participants in the control group (unchanged environment). This preliminary observation gives credence to the validity of our setup in inducing the target behavior of disorientation. We also observed that the patients with dementia expressed a considerable amount of disorientation. This was the case even though the environment remained unchanged, thereby indicating that moments of disorientation accruing to cognitive deficits can also be reliably observed among patients with dementia in our setup. The variations in motion and physiological data observed among our pilot participants illustrates that we are able to adequately obtain the relevant motion and physiological parameters within our approach. As building situation-adaptive devices for supporting navigation among people with dementia relies on the adequate detection of disorientation [6], our setup therefore provides a promising approach to acquiring the needed data.

We acknowledge that there are multiple alternatives to using the GRAIL system for the goal of investigating the effect of spatial disorientation on gait in a lab environment. An alternative to projecting the VR to a screen would be to use a VR headset, like the Oculus Rift (Facebook Inc) or the Samsung Gear VR (Samsung Electronics). A main advantage of the GRAIL over the Oculus Rift and Samsung Gear VR lies in the relative ease of use, safety, and closeness to reality for our study sample, which consists of cognitively healthy older adults and people with dementia who may not be as apt as younger participants at learning to navigate with a VR headset. To participate in the navigation task on the GRAIL, the participant walks freely on the treadmill without any headset or hand gear and observes the virtual environment as one normally would in a real-life situation. In contrast, with the Oculus Rift and Samsung Gear VR, a remarkable difference to the real-world situation is observed the moment one has to put on a headset and navigate with the use of controllers. Additionally, such devices can quickly cause simulator sickness, especially if participants have to move through the environment and when synchronization of head movement and VR movement is not optimal [46]. In this study, simulator sickness was only reported by 1 of the participants so far. Moreover, virtual reality games [such as serious games for dementia care (SGDC)] in which participants interact with the virtual environment using their body movements in the absence of controllers have also reported

positive effects on patients' mental and physical health [47]; these include improvement in memory and motivation, through the performance of mental and physical activities that are carried out while using various physical motions.

A major experience reported by a few of the participants, which also serves as a limitation to the current study, was the fact that control over left/right movement felt somewhat challenging. Possible alternatives to the current direction control mechanism used in this study include using a game controller to choose directions [19], or leaning in the respective direction [48]. However, we expect that this would be perceived as even more unrealistic in terms of how well the experimental setup models the real-world situation in which we are very much interested, as this would require movements that are not representative of real-world locomotion. For instance, Kizony et al [19] acknowledged that the combination of the joystick and treadmill may have posed an additional challenge to the older adults, although no discomfort was reported. This could have been further confirmed if participants were asked about their experience using structured measures such as the IPQ and usability questionnaires as employed in this study. Nonetheless, usability and immersion in VR studies remains an open topic. Further studies replicating the setup employed in this study could serve to provide further testing and data for evaluating this setup. Omnidirectional treadmills [49,50] (that allow free movement in all directions) could be a promising alternative for this purpose. However, such systems are still research prototypes and are not commercially available.

In conclusion, our GRAIL setup allows us to collect a rich dataset: most prominently, the motion capturing system can be used to assess spatio-temporal gait parameters (eg, walking speed). Results from this feasibility study suggest that our setup is sufficient in investigating sensor-derived features of spatial disorientation, even in patients with dementia, in a safe and controlled environment that is comparable to the real world. This is evident in the fact that we observed a considerable amount of disorientation among our participants, which also corresponded to the sensor data and questionnaire responses. Based on the full data set, in the future, we will focus on training machine learning classifier models for discriminating instances of disorientation from moments of orientation, with the aim of coming up with a relevant feature set for developing an interventional device. This classifier subsequently will be tested with real-world data (ie, perform transfer of learning). This feasibility study encourages further investigations on the possibility of more robust real-time detection of spatial disorientation, with the aim of coming up with adequate situation-aware interventions.

XSL•FO

RenderX

Authors' Contributions

COA and SL contributed to the conception of the work, acquisition and analysis of the data, and wrote substantial parts of the manuscript. COA reviewed and adapted the annotation scheme used for the study. TK, SJT, SL, and COA provided substantial contributions to the conception and design of the work and the acquisition, analysis, and interpretation of data for the work. TK, SJT, SL, and COA participated in drafting the work, revising it critically for important intellectual content, and approved the version to be published; these authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

None declared.

References

1. Kaye J, Matteck N, Hayes T, Austin D, Dodge H. P3-089: Community life withdrawal with MCI progression. 2012 Jul 01 Presented at: Alzheimer 's Association International Conference; July 14 - 19, 2012; Vancouver, British Columbia p. P483-P483. [doi: 10.1016/j.jalz.2012.05.1308]

2. Wettstein M, Wahl H, Shoval N, Oswald F, Voss E, Seidl U, et al. Out-of-home behavior and cognitive impairment in older adults: findings of the SenTra Project. J Appl Gerontol 2015 Feb;34(1):3-25. [doi: 10.1177/0733464812459373] [Medline: 25548086]

3. Glass TA, de Leon CM, Marottoli RA, Berkman LF. Population based study of social and productive activities as predictors of survival among elderly Americans. BMJ 1999 Aug 21;319(7208):478-483 [FREE Full text] [doi: 10.1136/bmj.319.7208.478] [Medline: 10454399]

4. Hughes TF, Flatt JD, Fu B, Chang CH, Ganguli M. Engagement in social activities and progression from mild to severe cognitive impairment: the MYHAT study. Int Psychogeriatr 2013 Apr;25(4):587-595 [FREE Full text] [doi: 10.1017/S1041610212002086] [Medline: 23257280]

5. Teipel S, Babiloni C, Hoey J, Kaye J, Kirste T, Burmeister OK. Information and communication technology solutions for outdoor navigation in dementia. Alzheimers Dement 2016 Jun;12(6):695-707 [FREE Full text] [doi: 10.1016/j.jalz.2015.11.003] [Medline: 26776761]

6. Schaat S, Koldrack P, Yordanova K, Kirste T, Teipel S. Real-Time Detection of Spatial Disorientation in Persons with Mild Cognitive Impairment and Dementia. Gerontology 2020;66(1):85-94. [doi: 10.1159/000500971] [Medline: 31362286]

7. Buracchio T, Dodge HH, Howieson D, Wasserman D, Kaye J. The trajectory of gait speed preceding mild cognitive impairment. Arch Neurol 2010 Aug;67(8):980-986 [FREE Full text] [doi: 10.1001/archneurol.2010.159] [Medline: 20697049]

8. Allali G, Annweiler C, Blumen HM, Callisaya ML, De Cock A, Kressig RW, et al. Gait phenotype from mild cognitive impairment to moderate dementia: results from the GOOD initiative. Eur J Neurol 2016 Mar;23(3):527-541 [FREE Full text] [doi: 10.1111/ene.12882] [Medline: 26662508]

9. Jayakody O, Breslin M, Srikanth VK, Callisaya ML. Gait Characteristics and Cognitive Decline: A Longitudinal Population-Based Study. J Alzheimers Dis 2019;71(s1):S5-S14. [doi: 10.3233/JAD-181157] [Medline: 30958358]

10. Callisaya ML, Blizzard L, Schmidt MD, Martin KL, McGinley JL, Sanders LM, et al. Gait, gait variability and the risk of multiple incident falls in older people: a population-based study. Age Ageing 2011 Jul;40(4):481-487. [doi: 10.1093/ageing/afr055] [Medline: 21628390]

11. Westmoreland D, Krell RW, Self BP. Physiological responses to the Coriolis illusion: effects of head position and vision. Aviat Space Environ Med 2007 Oct;78(10):985-989. [doi: 10.3357/asem.2010.2007] [Medline: 17955949]

12. Tamura A, Iwamoto T, Ozaki H, Kimura M, Tsujimoto Y, Wada Y. Wrist-Worn Electrodermal Activity as a Novel Neurophysiological Biomarker of Autonomic Symptoms in Spatial Disorientation. Front. Neurol 2018 Dec 4;9. [doi: 10.3389/fneur.2018.01056]

13. Cushman LA, Stein K, Duffy CJ. Detecting navigational deficits in cognitive aging and Alzheimer disease using virtual reality. Neurology 2008 Sep 16;71(12):888-895 [FREE Full text] [doi: 10.1212/01.wnl.0000326262.67613.fe] [Medline: 18794491]

14. Craig M, Wolbers T, Harris MA, Hauff P, Della Sala S, Dewar M. Comparable rest-related promotion of spatial memory consolidation in younger and older adults. Neurobiol Aging 2016 Dec;48:143-152 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2016.08.007] [Medline: 27689512]

15. McGuinness B, Barrett SL, Craig D, Lawson J, Passmore AP. Attention deficits in Alzheimer's disease and vascular dementia. J Neurol Neurosurg Psychiatry 2010 Feb;81(2):157-159. [doi: 10.1136/jnnp.2008.164483] [Medline: 20145028]

16. Silveri MC, Reali G, Jenner C, Puopolo M. Attention and memory in the preclinical stage of dementia. J Geriatr Psychiatry Neurol 2007 Jun;20(2):67-75. [doi: 10.1177/0891988706297469] [Medline: 17548775]

17. Mirahadi SS, Khatoonabadi SA, Fekar Gharamaleki F. A Review of Divided Attention Dysfunction in Alzheimer's Disease. Middle East J Rehabil Health Stud 2018 Jun 26;5(3). [doi: 10.5812/mejrh.64738]

18. Nebes RD, Brady CB. Focused and Divided Attention in Alzheimer's Disease. Cortex 1989 Jun;25(2):305-315. [doi: 10.1016/s0010-9452(89)80045-0]

19.   Kizony R, Zeilig G, Krasovsky T, Bondi M, Weiss PL, Kodesh E, et al. Using virtual reality simulation to study navigation in a complex environment as a functional-cognitive task; A pilot study. J Vestib Res 2017;27(1):39-47. [doi: 10.3233/VES-170605] [Medline: 28387691]

20.   Ijaz K, Ahmadpour N, Naismith SL, Calvo RA. An Immersive Virtual Reality Platform for Assessing Spatial Navigation Memory in Predementia Screening: Feasibility and Usability Study. JMIR Ment Health 2019 Sep 03;6(9):e13887 [FREE Full text] [doi: 10.2196/13887] [Medline: 31482851]

21.   Moreno A, Wall KJ, Thangavelu K, Craven L, Ward E, Dissanayaka NN. A systematic review of the use of virtual reality and its effects on cognition in individuals with neurocognitive disorders. Alzheimers Dement (N Y) 2019;5:834-850 [FREE Full text] [doi: 10.1016/j.trci.2019.09.016] [Medline: 31799368]

22.   Kober SE, Wood G, Hofer D, Kreuzig W, Kiefer M, Neuper C. Virtual reality in neurologic rehabilitation of spatial disorientation. J Neuroeng Rehabil 2013 Feb 08;10:17 [FREE Full text] [doi: 10.1186/1743-0003-10-17] [Medline: 23394289]

23.   van den Bogert AJ, Geijtenbeek T, Even-Zohar O, Steenbrink F, Hardin EC. A real-time system for biomechanical analysis of human movement and muscle function. Med Biol Eng Comput 2013 Oct;51(10):1069-1077 [FREE Full text] [doi: 10.1007/s11517-013-1076-z] [Medline: 23884905]

24.   de Rooij IJM, van de Port IGL, Visser-Meily JMA, Meijer JG. Virtual reality gait training versus non-virtual reality gait training for improving participation in subacute stroke survivors: study protocol of the ViRTAS randomized controlled trial. Trials 2019 Jan 29;20(1):89 [FREE Full text] [doi: 10.1186/s13063-018-3165-7] [Medline: 30696491]

25.   de Kruijff LGM, Prins M, van der Krans A, Hoencamp R, van der Wurff P. Combat-related foot injuries: impact on gait and functional outcome. J R Army Med Corps 2018 Sep;164(5):322-327. [doi: 10.1136/jramc-2017-000870] [Medline: 29436477]

26.   Biffi E, Costantini C, Ceccarelli SB, Cesareo A, Marzocchi GM, Nobile M, et al. Gait Pattern and Motor Performance During Discrete Gait Perturbation in Children With Autism Spectrum Disorders. Front Psychol 2018;9:2530 [FREE Full text] [doi: 10.3389/fpsyg.2018.02530] [Medline: 30618953]

27.   Behrens M, Mau-Moeller A, Lischke A, Katlun F, Gube M, Zschorlich V, et al. Mental Fatigue Increases Gait Variability During Dual-task Walking in Old Adults. J Gerontol A Biol Sci Med Sci 2018 May 09;73(6):792-797. [doi: 10.1093/gerona/glx210] [Medline: 29077783]

28.   Knerr T. OSM2World. 2019. URL: http://osm2world.org/ [accessed 2020-06-07]

29.   Sloot LH, van der Krogt MM, Harlaar J. Self-paced versus fixed speed treadmill walking. Gait Posture 2014;39(1):478-484 [FREE Full text] [doi: 10.1016/j.gaitpost.2013.08.022] [Medline: 24055003]

30.   Yordanova K, Koldrack P, Heine C, Henkel R, Martin M, Teipel S, et al. Situation Model for Situation-Aware Assistance of Dementia Patients in Outdoor Mobility. J Alzheimers Dis 2017;60(4):1461-1476. [doi: 10.3233/JAD-170105] [Medline: 29060937]

31.   Härtel S, Gnam J, Löffler S, Bös K. Estimation of energy expenditure using accelerometers and activity-based energy models—validation of a new device. Eur Rev Aging Phys Act 2010 Dec 2;8(2):109-114. [doi: 10.1007/s11556-010-0074-5]

32.   Kapp D, Schaaff K, Ottenbacher J, Heuer S, Bachis S. Isolating the effects of emotional stimuli in EDA measurements: White paper 2017. URL: https://www.movisens.com/de/produkte/eda-sensor/ [accessed 2020-07-07]

33.   Bachis S, Ottenbacher J. Validity of ecgMove R-peak detection: White paper 2017. URL: https://www.movisens.com/de/produkte/ekg-sensor/ [accessed 2020-07-07]

34.   Schubert T, Friedmann F, Regenbrecht H. The Experience of Presence: Factor Analytic Insights. Presence: Teleoperators and Virtual Environments 2001 Jun;10(3):266-281. [doi: 10.1162/105474601300343603]

35.   Urbaniak G, Plous S. Research Randomizer (Version 4.0). 2013. URL: https://www.randomizer.org/ [accessed 2018-12-12]

36.   Wittenburg P, Brugman H, Russel A, Klassmann A, Sloetjes H. ELAN: a professional framework for multimodality research. 2006 Presented at: 5th International Conference on Language Resources and Evaluation (LREC 2006); 2006; Genoa p. 1556-1559. [doi: 10.1093/oxfordhb/9780199571932.013.019]

37.   Cohen J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 1960 Apr 01;20(1):37-46. [doi: 10.1177/001316446002000104]

38.   da Costa RQM, Pompeu JE, de Mello DD, Moretto E, Rodrigues FZ, Dos Santos MD, et al. Two new virtual reality tasks for the assessment of spatial orientation Preliminary results of tolerability, sense of presence and usability. Dement Neuropsychol 2018;12(2):196-204 [FREE Full text] [doi: 10.1590/1980-57642018dn12-020013] [Medline: 29988338]

39.   Delahaye M, Lemoine P, Cartwright S, Deuring G, Beck J, Pflueger M, et al. Learning aptitude, spatial orientation and cognitive flexibility tested in a virtual labyrinth after virtual stress induction. BMC Psychol 2015;3(1):22 [FREE Full text] [doi: 10.1186/s40359-015-0080-5] [Medline: 26146558]

40.   Kimura K, Reichert JF, Olson A, Pouya OR, Wang X, Moussavi Z, et al. Orientation in Virtual Reality Does Not Fully Measure Up to the Real-World. Sci Rep 2017 Dec 22;7(1):18109 [FREE Full text] [doi: 10.1038/s41598-017-18289-8] [Medline: 29273759]

41.   Burgess N, Trinkler I, King J, Kennedy A, Cipolotti L. Impaired allocentric spatial memory underlying topographical disorientation. Rev Neurosci 2006;17(1-2):239-251. [doi: 10.1515/revneuro.2006.17.1-2.239] [Medline: 16703955]

42.   Tichon J, Banks J. Virtual reality exposure therapy: 150-degree screen to desktop PC. Cyberpsychol Behav 2006 Aug;9(4):480-489. [doi: 10.1089/cpb.2006.9.480] [Medline: 16901251]

43.   Mendez MF, Tomsak RL, Remler B. Disorders of the visual system in Alzheimer's disease. J Clin Neuroophthalmol 1990 Mar;10(1):62-69. [Medline: 2139054]

44.   Kimoto A, Iseki E, Ota K, Murayama N, Sato K, Ogura N, et al. Differences in responses to the Rorschach test between patients with dementia with Lewy bodies and Alzheimer's disease -from the perspective of visuoperceptual impairment. Psychiatry Res 2017 Nov;257:456-461. [doi: 10.1016/j.psychres.2017.08.038] [Medline: 28837937]

45.   Cohen J, Verghese J. Gait and dementia. Handb Clin Neurol 2019;167:419-427. [doi: 10.1016/b978-0-12-804766-8.00022-4] [Medline: 31753146]

46.   Nichols S, Patel H. Health and safety implications of virtual reality: a review of empirical evidence. Applied Ergonomics 2002 May;33(3):251-271. [doi: 10.1016/s0003-6870(02)00020-0]

47.   Breton Z, Zapirain BG, Zorrilla AM. KiMentia: Kinect based tool to help cognitive stimulation for individuals with dementia. : IEEE; 2012 Presented at: 2012 IEEE 14th International Conference on e-Health Networking, Applications and Services; 10-13 October 2012; Beijing, China. [doi: 10.1109/healthcom.2012.6379430]

48.   Tregillus S, Zayer M, Folmer E. Handsfree Omnidirectional VR Navigation using Head Tilt. New York, NY: The Association for Computing Machinery; 2017 Presented at: 2017 CHI Conference on Human Factors in Computing Systems (CHI '17); May 2017; New York p. 4063-4066. [doi: 10.1145/3025453.3025521]

49.   Iwata H. The Torus Treadmill: realizing locomotion in VEs. IEEE Comput. Grap. Appl 1999 Nov 01;19(6):30-35. [doi: 10.1109/38.799737]

50.   Souman JL, Giordano PR, Schwaiger M, Frissen I, Thümmel T, Ulbrich H, et al. CyberWalk. ACM Trans. Appl. Percept 2011 Nov;8(4):1-22. [doi: 10.1145/2043603.2043607]

Abbreviations

AD: Alzheimer disease
ATD: assistive technology device
BPM: beats per minute
ECG: electrocardiography
EDA: electrodermal activity
GRAIL: Gait Real-time Analysis Interactive Lab
IPQ: igroup Presence Questionnaire
MCI: mild cognitive impairment
MMSE: Mini–Mental State Examination
OSM: OpenStreetMap
SiNDeM: Situation-Aware Navigation Assistance for Dementia Patients using Causal Behavior Models
VR: virtual reality

XSL•FO
RenderX

Original Paper

Use of a Low-Cost Portable 3D Virtual Reality Simulator for Psychomotor Skill Training in Minimally Invasive Surgery: Task Metrics and Score Validity

Fernando Alvarez-Lopez[1,2], MD; Marcelo Fabián Maina[3], PhD; Fernando Arango[1], MD; Francesc Saigí-Rubió[2], PhD

[1]Faculty of Health Sciences, Universidad de Manizales, Manizales, Colombia

[2]Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain

[3]Faculty of Psychology and Education Sciences, Universitat Oberta de Catalunya, Barcelona, Spain

Corresponding Author:
Francesc Saigí-Rubió, PhD
Faculty of Health Sciences
Universitat Oberta de Catalunya
Avinguda Tibidabo, 39-43
Barcelona, 08035
Spain
Phone: 34 933 263 622
Email: fsaigi@uoc.edu

Abstract

Background: The high cost and low availability of virtual reality simulators in surgical specialty training programs in low- and middle-income countries make it necessary to develop and obtain sources of validity for new models of low-cost portable simulators that enable ubiquitous learning of psychomotor skills in minimally invasive surgery.

Objective: The aim of this study was to obtain validity evidence for relationships to other variables, internal structure, and consequences of testing for the task scores of a new low-cost portable simulator mediated by gestures for learning basic psychomotor skills in minimally invasive surgery. This new simulator is called SIMISGEST-VR (Simulator of Minimally Invasive Surgery mediated by Gestures - Virtual Reality).

Methods: In this prospective observational validity study, the authors looked for multiple sources of evidence (known group construct validity, prior videogaming experience, internal structure, test-retest reliability, and consequences of testing) for the proposed SIMISGEST-VR tasks. Undergraduate students (n=100, reference group), surgical residents (n=20), and experts in minimally invasive surgery (n=28) took part in the study. After answering a demographic questionnaire and watching a video of the tasks to be performed, they individually repeated each task 10 times with each hand. The simulator provided concurrent, immediate, and terminal feedback and obtained the task metrics (time and score). From the reference group, 29 undergraduate students were randomly selected to perform the tasks 6 months later in order to determine test-retest reliability.

Results: Evidence from multiple sources, including strong intrarater reliability and internal consistency, considerable evidence for the hypothesized consequences of testing, and partial confirmation for relations to other variables, supports the validity of the scores and the metrics used to train and teach basic psychomotor skills for minimally invasive surgery via a new low-cost portable simulator that utilizes interaction technology mediated by gestures.

Conclusions: The results obtained provided multiple sources of evidence to validate SIMISGEST-VR tasks aimed at training novices with no prior experience and enabling them to learn basic psychomotor skills for minimally invasive surgery.

Introduction

Background

The advent of minimally invasive surgery in the mid-1980s [1] led to an increase in the number of iatrogenic bile duct injuries when many surgeons worldwide switched from open surgery to minimally invasive surgery without any prior training [2-8]. As a result, simulation has since become a valuable tool for learning motor skills for minimally invasive surgery. Many studies have demonstrated that simulation is a useful tool for learning motor skills for minimally invasive surgery and that learned skills can be transferred to the operating theatre [9-19].

The first virtual reality (VR) simulator for minimally invasive surgery training was MIST-VR (Minimally Invasive Surgery Training - Virtual Reality) [20]. Evidence for construct validity was established in 1998 [21], and evidence for predictive validity was obtained in 2002 [9,22]. Subsequently, evidence for concurrent validity was also demonstrated [23-25].

Recent years have seen the development of low-cost gesture-based touchless devices that can interact with 3D virtual environments, among them Kinect (Microsoft Corp), the Leap Motion Controller (Leap Motion Inc), and the Myo armband (Thalmic Labs) [26].

The Leap Motion Controller was launched in May 2012. It is based on the principle of infrared optical tracking, which detects the positions of fine objects such as fingertips or pen tips in a Cartesian plane; its interaction zone is an inverted cone of approximately 0.23 m^3, and it has a motion detection range between 20 mm and 600 mm [27,28]. It measures 76 mm $\times$ 30 mm $\times$ 13 mm and weighs 45 g. It has 3 infrared emitters and 2 infrared cameras that capture the movements generated within the interaction zone [29,30]. The manufacturer reports an accuracy of 0.01 mm for fingertip detection, although one independent study showed an accuracy of 0.7 mm [31]. Although the Leap Motion Controller is designed mainly to detect hand motions, it can track objects such as pencils and laparoscopic surgical forceps [32-34].

The Leap Motion Controller has been used as a tool for the manipulation of medical images in the fields of interventional radiology and image-guided surgery or when there is a risk of contamination through contact (autopsy rooms, for example), for touchless control (operating theatre lights and tables) and for simulation (minimally invasive surgery and robotic surgery). Various authors have used the Leap Motion Controller to develop simulators that track hand or instrument movements [26,32-39]. A paper by Lahanas [35] describes using Leap Motion Controller to simulate 3 tasks: camera navigation, instrument navigation, and bimanual operation; 28 expert surgeons and 21 reference individuals took part in the study. The experts significantly outperformed novices in all assessment metrics for instrument navigation and bimanual operation.

Simulators for learning skills for minimally invasive surgery can be classified into 3 types: traditional box trainers, augmented reality simulators (hybrid), and VR simulators [40,41]. Simulation has become a valuable tool for learning basic motor skills in surgery, but access to simulators remains problematic, especially in low- and middle-income countries, because of their high cost. Consequently, that makes it necessary to develop and validate the metrics and scores of low-cost portable simulators [42-44].

The aim of this study was to evaluate a simulation instrument, SIMISGEST-VR (Simulator of Minimally Invasive Surgery mediated by Gestures - Virtual Reality), and to document the sources of validity evidence for task scores, relations to other variables, internal structure, consequences of testing, and response process.

Hypotheses

To that end, 3 hypotheses were formulated:

Hypothesis 1: Validity Evidence for Relations to Other Variables

The first hypothesis aims to demonstrate that the test scores discriminate between a reference group (no prior experience), surgical residents (less experienced), and surgeons (more experienced), showing that the experts already have the basic psychomotor skills being measured, and similarly, that videogaming experience is correlated with better performance in simulator tasks, regardless of the level of training and experience.

Hypothesis 2: Evidence for Internal Structure

The intrarater test-retest assumes that, if a reference individual is not exposed to simulators in the period of time between the 2 complete simulator exercises, there will be no significant differences in performance between the first and second exercises.

Hypothesis 3: Evidence for Consequences of Testing

Regarding evidence for consequences of testing, the reference group learning will be demonstrated by improvements in the metrics and the final score when comparing the first and the tenth attempt in each task.

Methods

Study Design

This was a prospective observational validity study. The current unified standard considers that all validity is construct validity and, as such, requires evidence from 5 sources [45-52].

Content evidence includes a description of the steps taken to ensure that test content reflects, in a relevant way, the construct or characteristic being measured. The results obtained from the survey assessing fidelity to the criterion and content-related validity evidence for SIMISGEST-VR showed that all 30 participants felt that most aspects of the simulator were adequately realistic and that it could be used as a tool for teaching basic psychomotor skills in laparoscopic surgery (Likert score: range 4.07-4.73). The sources of content-related validity evidence showed that our simulator was a reliable training tool and that the tasks enabled learning of the basic psychomotor skills required in minimally invasive surgery (Likert score: range 4.28-4.67) [53].

Evidence for *relations to other variables* refers to the statistical association between the test scores and other characteristics or external measures that have theoretical relations, such as level of training, level of experience, prior videogaming experience, and scores for other already validated instruments. One of the most common correlations is *known group construct validity* (ie, the correlation between performance scores and level of training and experience) [54]. Relations may be positive (convergent or predictive) or negative (divergent or discriminant) depending on the constructs being measured [55]. This study explored the relations between performance scores and the level of training, experience, and prior videogaming experience.

Evidence for *internal structure* includes data that evaluate the relations between the individual items of the assessment, and how they correlate to the construct. It includes measures of reliability, reproducibility, and factor analysis. Reliability is a necessary but insufficient condition for validity [56]. Intrarater reliability was obtained using the test-retest method, which assesses the stability of responses over time [57]. Test-retest reliability was explored through blinded rerating after an interval of 6 months in the reference group. The randomly selected participants were asked whether they had had additional experience of using simulators during that period of time [56]. The answer was "no" in all cases. The data produced by this second test were not taken into account in the evidence for the construct validity study. Worster and Haines [58] noted that there was no published recommendation for the proportion of data that should be checked but that 10% was common. In this study, 29% of the reference individuals were included in the test-retest study. The demonstration of reliability is mandatory before an evaluation can be shown to be valid [54].

Evidence for *consequences* refers to the impact, benefit, or danger of assessment itself and the resulting decisions and actions. Yet, simply demonstrating consequences, even significant and impressive ones, does not constitute validity evidence unless investigators explicitly demonstrate that these consequences have an impact on score interpretation (validity) [46,55]. Evidence for consequences falls within a spectrum between high-stake examinations, licensing examinations, or low-stake examinations such a self-assessment used for formative feedback alone [54]. In our case, we hoped to obtain evidence to demonstrate that the reference group had managed to achieve the learning curve.

Evidence for *response process* includes theoretical and empirical analyses evaluating the extent to which the assessors' and respondents' responses are aligned to the construct. It includes an evaluation of safety, of quality control, and of the actors' thoughts and actions during the assessment. The response process also includes the accuracy of data collection and entry into the database [54]. This type of evidence can be difficult to demonstrate because data are often qualitative [55].

Participants and Simulator Test Methodology

Participating in this study were minimally invasive surgery expert surgeons (n=28) in a range of surgical specialties, each who had performed more than 100 procedures, surgical residents (n=20) in a range of surgical specialties from the University of

Caldas (in Manizales, Colombia), each who had performed fewer than 50 procedures (basic training: n=15; advanced training: n=5), and medical undergraduate students (n=100) from the University of Caldas and the University of Manizales who had no experience performing minimally invasive surgical procedures. The expert surgeons worked in the following specialties: general surgery 8 (28.5%), pediatric surgery 5 (17.8%), neurosurgery 4 (14.2%), colorectal surgery 3 (10.7%), orthopedic surgery 3 (10.7%), gynecological surgery 2 (7.1%), urological surgery 1 (3.5%), thoracic surgery 1 (3.5%), and vascular surgery 1 (3.5%).

All participants completed a questionnaire providing demographic data (Multimedia Appendix 1) and information about the dominant hand, level of training, levels of minimally invasive surgery skills, prior training with simulators, and experience with videogaming or VR devices.

After the instructor had given basic instructions about using the simulator and had shown a video of each task to be performed, the study participants performed 10 repetitions of tasks 1, 2, 4, 5, and 6 with each hand. Task 3 was repeated 10 times because both hands were considered dominant. The instructor did not give additional feedback, but the simulator did provide concurrent feedback (visual and auditory feedback while performing each task), immediate feedback (displaying the results in terms of time, accuracy and errors at the end of each task), and terminal feedback (performance curve and final score). The participants were able to watch the demonstration videos again at any time. For the test-retest reliability study, 29 participants were randomly selected from the reference group. They repeated the entire exercise 6 months after the first exercise; none were exposed to any type of simulator during that period of time. One of the authors (FAL) supervised and photographically documented each exercise.

SIMISGEST-VR

SIMISGEST-VR was developed using design-based research [59-63]. A previously published article [53] describes in detail the development of the device and a study assessing fidelity to the criterion and content-related validity evidence.

Virtual Environment

The virtual environment consisted of the following modules: registration to collect users' demographic data and a tutorial to show demonstration videos of the tasks to be performed.

SIMISGEST-VR supports 6 tasks, each of which corresponds to a surgical equivalent (Table 1; Figure 1). The tasks were adapted from MIST-VR (Mentice Inc) [20,64,65]. MIST-VR is the simulator on which the highest number of validation studies have been conducted, and they have demonstrated, on multiple occasions, that the skills that are learned can be transferred to the operating theatre [9,21,66-73].

Except for task 3, all tasks had the option of configuring the dominant hand during the exercise; task 3 required the simultaneous use of both hands and therefore both played a dominant function. Given its level of difficulty, this task was performed last in all cases. The online virtual environment ran on Windows (Microsoft Inc) and MacOS (Apple Inc) platforms.

Table 1. Description of the tasks and their surgical equivalents. Adapted from Sutton et al [20].

Task number	Task name	Description	Surgical equivalent	Learning objective
Task 1	Grip and placement	Take the sphere with one hand and move it to a new location within the workspace	Gripping and retraction of tissue to a given position, placement of clips and hemostasis, use of extractor bags	Visual-spatial perception and eye-hand coordination
Task 2	Transfer and placement of an object	Take the sphere, transfer it to another instrument and place it inside a hollow cylinder	Transfer of a needle between a clamp and a needle holder	Visual-spatial perception, eye-hand coordination, and use of both hands in a complementary manner
Task 3	Cross	Instruments travel along a surface in a 3D cylinder	Small intestine exploration	Coordinated use of both the dominant and nondominant hands and ambidexterity
Task 4	Removal and reinsertion of instruments	Removal of instruments from the operative site and reinsertion	One instrument stabilizes one organ while the other is removed from the field and reintroduced	Visual-spatial perception, use of both hands in a complementary manner, and depth perception
Task 5	Diathermy	Cauterize a series of targets located in a fixed sphere	Cauterize a bleeding blood vessel	Visual-spatial perception, time of diathermy, and accuracy of movements
Task 6	Target manipulation and diathermy	Take the sphere with the instrument and place it inside a virtual space represented by a cube and cauterize a series of targets with the other hand.	Present and set a target to cauterize	Visual-spatial perception, time of diathermy, and accuracy of movements

XSL·FO
RenderX

Figure 1. Task screenshots.
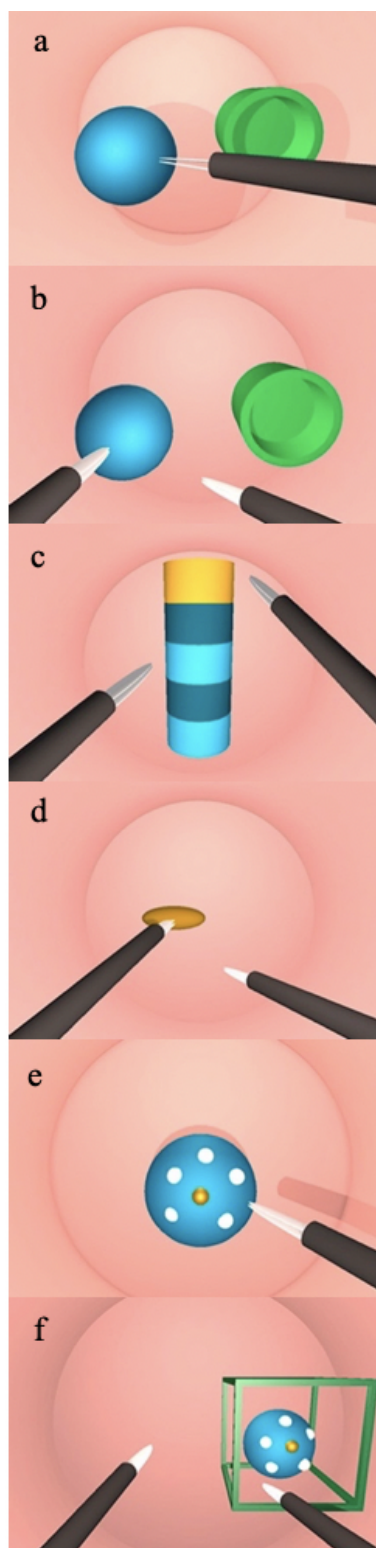
Metrics

The metrics were established using 5 parameters: time (velocity), efficiency of movement for the right and left hands [21,74], economy of diathermy, error and accuracy (penalty) [75,76], and final score.

Feedback

Feedback is essential [77]. Training on a simulator should have 3 purposes: to improve performance, to make performance consistent, and to reduce the number of errors [78]. The haptic sensation and concurrent feedback were simulated using sound signals, color changes in the objects, and movement of the object when an undue collision occurred between the different components of the environment or when an error occurred during the task (concurrent feedback). For SIMISGEST-VR,

we adopted 3 types of feedback: concurrent, which was provided while the task is being performed; immediate, which was provided at the end of each task when the system provides information on the presence or absence of errors, efficiency, and the time taken; and terminal, which was provided at the end of each training session when the system provides a series of graphs and tables that show performance over time [79-83].The data generated by the program were stored on an SQL (structured query language) database engine integrated into the simulation software.

Hardware

Two laparoscopic forceps were used. In fact, we used simulated forceps made using 3D printers. These minimally invasive

surgery forceps did not need to be functional. The final device with all its components assembled is shown in Figure 2. Figure 2 shows the fixing pad (1) for the Leap Motion Controller and the mounting support devices (3) for the minimally invasive surgery laparoscopic forceps (2), which allow simulation of the fulcrum effect; the Leap Motion Controller (4), responsible for detecting the movements of the instruments; and the computer, which, by means of the software programs running on it, administers the virtual environment and the metrics, and provides feedback and the final performance score on the screen (5) where the 3D virtual environment is displayed.

To perform the test, a 13-inch MacBook Pro (Apple Inc) was used, which served as a screen, ran the 3D virtual environment, and stored metrics data.

Figure 2. Diagram of the artifact. MIS: minimally invasive surgery.

1. Support pad:
   - Length: 35 cm
   - Width: 25 cm
   - Height: 1.5 cm

2. Forceps of MIS:
   - Length of the stem: 33 cm

3. Forceps support device:
   - Height: 22.5 cm
   - Distance between the top of the devices: 16.5 cm

4. Leap Motion Controler TM®:
   - Length: 8 cm
   - Width: 3 cm
   - Height: 12 mm

5. Virtual environment
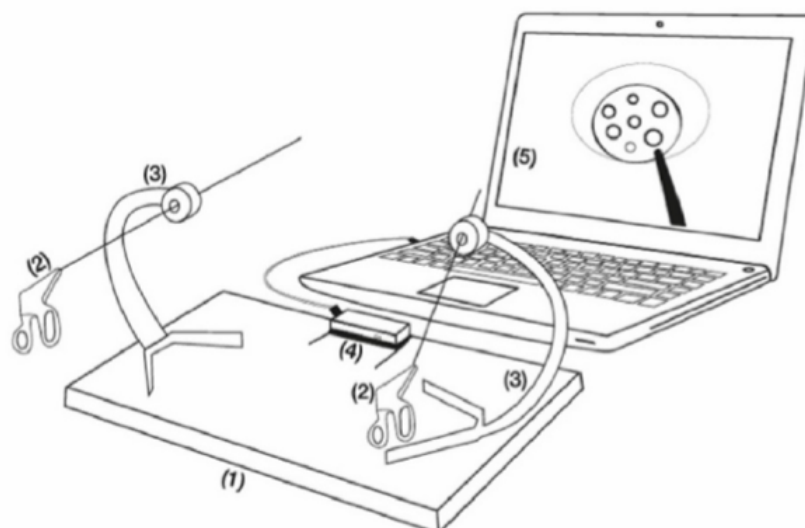
Data Analysis

Continuous data are presented in a frequency distribution table by mean and standard deviation. The Shapiro-Wilk test was used to assess normality. Categorical data are also presented in a frequency distribution table. Since the metrics data were not normally distributed, nonparametric tests were used to assess the hypotheses. Regarding hypothesis 1, the differences in the scores and time taken to perform the first trial in each task between novices and experts were compared using the Wilcoxon signed-rank test. Among the novices, the final scores of the tenth trial in each task were compared by prior videogaming experience using the Kolmogorov-Smirnov test. To assess hypothesis 2, internal consistency was calculated using Cronbach α. In addition, test-retest reliability was assessed by comparing the tenth trial in each task performed initially and repeated 6 months later using the Spearman correlation coefficient. To assess hypothesis 3, the scores and time taken in the first and last trials in each task were compared by level

of training using the Wilcoxon signed-rank test. In addition, excess diathermy in the first and last trials in tasks 5 and 6 was calculated by level of training using the Wilcoxon signed-rank test. $P<.05$ as level of statistical significance was established. Statistical analysis was performed using Stata (version 15.0; StataCorp LLC).

Results

Demographic Profile

Regarding prior experience with simulators, 35% (7/20) of the surgical residents and 36% (10/28) of the surgeons surveyed said they did not have any. Among the surgical residents, only 15% (3/20) had experience with VR simulators, and none had any experience with hybrid ones.

When videogaming experience was assessed, the low percentage of frequent gaming (daily or weekly) was striking: only 28% (28/100) in the reference group, 20% (4/20) among surgical residents, and 14% (4/28) among experts (Table 2).

Table 2. Demographic profile of study participants.

Variable	Reference group (n=100)	Surgical residents (n=20)	Surgeons (n=28)
Gender, n (%)			
Female	47 (47)	10 (50)	4 (15)
Male	53 (53)	10 (50)	24 (86)
Age	23.5 (0.28)	28.4 (0.54)	47 (2.12)
Dominant hand, n (%)			
Right	89 (89)	19 (95)	27 (96)
Left	11 (11)	1 (5)	1 (4)
Experience with simulators, n (%)			
Yes	1 (1)	13 (65)	18 (64)
No	99 (99)	7 (35)	10 (36)
Type of simulator, n (%)			
Not applicable	99 (99)	7 (35)	10 (25)
Virtual reality	1 (1)	3 (15)	7 (36)
Physical	0 (0)	10 (50)	10 (25)
Hybrid/augmented reality	0 (0)	0 (0)	1 (4)
Videogaming experience, n (%)			
Yes	72 (72)	15 (75)	14 (50)
No	28 (28)	5 (25)	14 (50)
Videogaming frequency, n (%)			
Not applicable	26 (26)	4 (20)	14 (50)
Daily	1 (1)	1 (5)	0 (0)
Weekly	27 (27)	3 (15)	4 (14)
Monthly	6 (6)	3 (15)	3 (11)
Occasionally	40 (40)	9 (45)	7 (25)
Minimally invasive surgery experience, n (%)			
None	37 (37)	3 (15)	0 (0)
Basic camera manipulation	63 (63)	6 (30)	0 (0)
Basic operator level	0 (0)	11 (55)	10 (36)
Intermediate operator level	0 (0)	0 (0)	10 (36)
Advanced operator level	0 (0)	0 (0)	8 (29)

Validity Hypothesis 1: Relations to Other Variables

To explore validity evidence for relations to other variables, we compared the SIMISGEST-VR test scores across experience levels (known group construct validity). No statistically significant differences were found in the scores of the first trial in each task between novices and experts; however, the times taken to perform tasks 3 (P=.006) and 6 (P=.02) were statistically significantly lower for experts compared to those of the reference group (Table 3). Performance in task 5 was better for novices who had prior videogaming experience (P=.01), as shown in Figure 3. When time was considered as a metric in task 3, a statistically significant difference (P=.006) was found between the reference group and the experts in performing the first trial (Figure 4).

Table 3. Trial 1 scores and time between novices and experts.

Metric and task	*P* value
Score	
Task 1	.58
Task 2	.13
Task 3	.33
Task 4	.18
Task 5	.77
Task 6	.27
Time	
Task 1	.53
Task 2	.34
Task 3	.006
Task 4	.26
Task 5	.28
Task 6	.02

Figure 3. Box plot of scores in task 5 performed by the reference group, by prior videogaming experience.
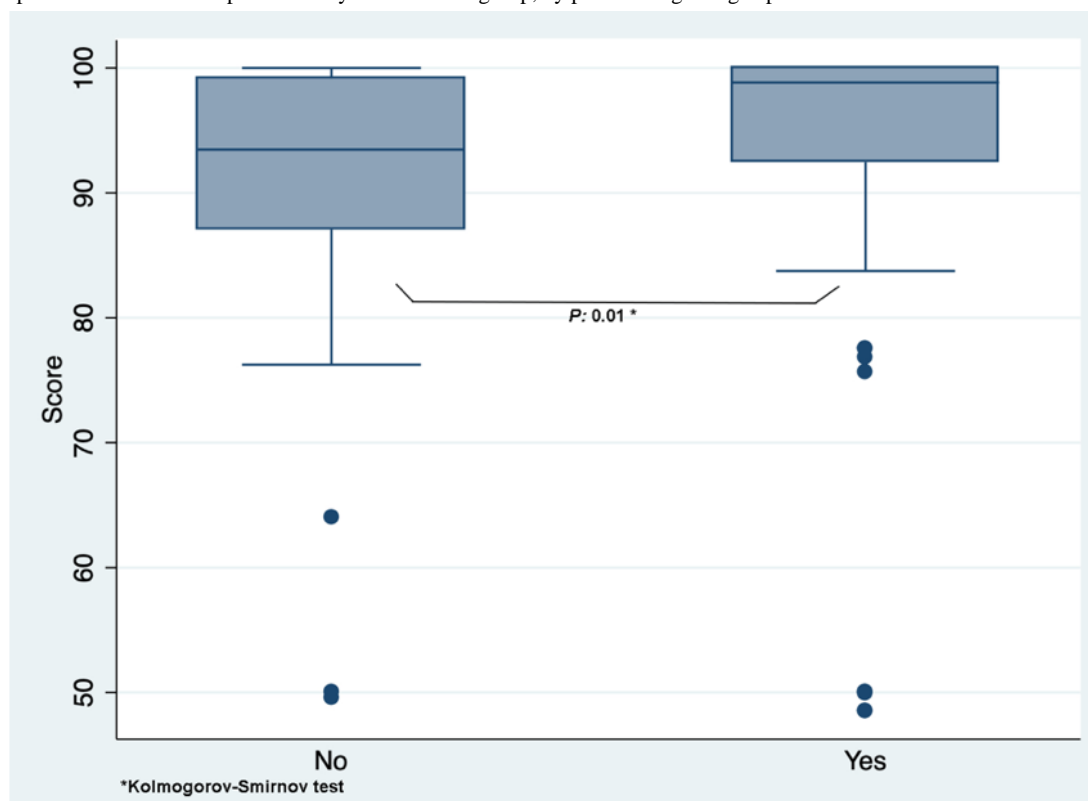
Figure 4. Box plot of times in task 3, by level of training.
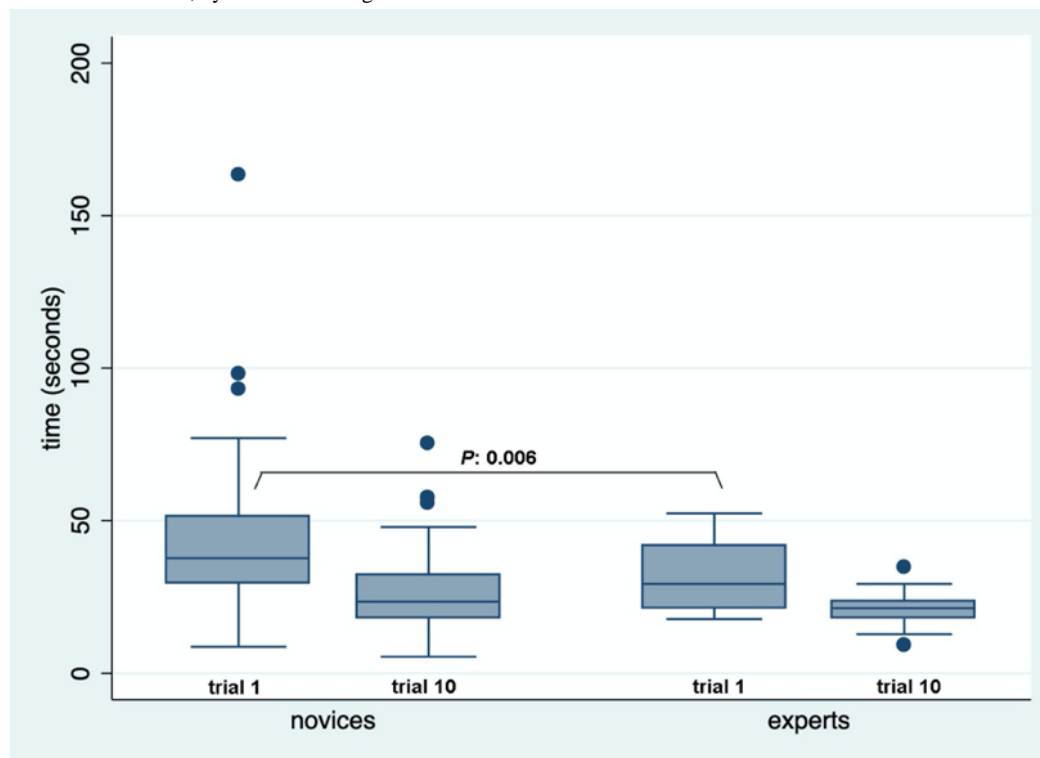
Validity Hypothesis 2: Internal Structure

The items demonstrated high internal consistency (Cronbach α=.81). Regarding the final scores in all the tasks, no statistically significant differences were found between the first exercises and those 6 months later for the randomly selected participants from the reference group (Table 4); when time was assessed as a metric, statistically significant differences were found for tasks 4 (trial 10: *P*=.048) and 6 (trial 10: *P*=.03). This demonstrates full evidence for the internal structure and test-retest reliability with respect to the score and partial evidence with respect to time as a metric (Table 4).

Table 4. Test-retest score for novices showing reliability between trials.

Metric, task, and trial comparison	Spearman correlation coefficient	*P* value
Score (initial vs 6 months later)		
Task 1		
Trial 1	0.200	.23
Trial 10	−0.294	.12
Task 2		
Trial 1	0.0846	.66
Trial 10	−0.256	.18
Task 3		
Trial 1	−0.036	.85
Trial 10	0.150	.44
Task 4		
Trial 1	0.120	.53
Trial 10	0.338	.07
Task 5		
Trial 1	0.341	.07
Trial 10	0.035	.85
Task 6		
Trial 1	−0.321	.09
Trial 10	−0.030	.88
Time (initial vs 6 months later)		
Task 1		
Trial 1	−0.005	.98
Trial 10	0.243	.20
Task 2		
Trial 1	−0.082	.67
Trial 10	−0.216	.26
Task 3		
Trial 1	0.121	.53
Trial 10	0.359	.06
Task 4		
Trial 1	−0.141	.46
Trial 10	0.370	.048
Task 5		
Trial 1	0.271	.16
Trial 10	0.330	.08
Task 6		
Trial 1	0.097	.62
Trial 10	0.412	.03

Validity Hypothesis 3: Consequences of Testing

Among the reference group, statistically significant differences were found in the scores and the time taken to perform each task between the first and tenth trials. Among the experts, statistically significant differences were found in the scores in tasks 1 ($P<.001$), 3 ($P=.03$), and 4 ($P=.01$), and in the time taken

to perform each task. These findings demonstrate a learning curve (Table 5).

Table 5. Score and time between trial 1 and trial 10, by level of training.

Metric and task	Comparison trials	Novices, *P* value	Experts, *P* value
Score			
Task 1	1 vs 10	<.001	<.001
Task 2	1 vs 10	.002	.052
Task 3	1 vs 10	.002	.03
Task 4	1 vs 10	<.001	.01
Task 5	1 vs 10	<.001	.31
Task 6	1 vs 10	.003	.63
Time			
Task 1	1 vs 10	<.001	<.001
Task 2	1 vs 10	<.001	<.001
Task 3	1 vs 10	<.001	<.001
Task 4	1 vs 10	<.001	<.001
Task 5	1 vs 10	<.001	<.001
Task 6	1 vs 10	<.001	<.001

In task 5, the reference group made statistically significantly fewer excess diathermy errors in the tenth trial than they did in the first trial (P=.003), which is evidence of a learning curve (Table 6).

Table 6. Excess diathermy errors when doing trials 1 and 10 in tasks 5 and 6, by level of training. Wilcoxon signed-rank test.

Task and group	Trial 1, mean (95% CI)	Trial 10, mean (95% CI)	*P* value
Task 5			
Novices	0.205 (0.127, 0.282)	0.090 (0.031, 0.148)	.003
Experts	0.125 (0.039, 0.210)	0.089 (0.013, 0.164)	.56
Task 6			
Novices	0.200 (0.105, 0.294)	0.205 (0.120, 0.289)	.90
Experts	0.250 (0.049, 0.450)	0.107 (0.010, 0.203)	.40

Response Process Validity

Study participants had the opportunity to observe each task in advance by watching a video, and they received basic instruction. The only feedback the participants received was from the simulator; they did not receive any other type of feedback from the instructor. Each of the 177 tests performed (148 initial tests and 29 test-retests) was supervised by the same person (FAL). Photographic documentation of every person performing the tasks was obtained. The final performance scores were defined in advance by using the formula described in another study [53]. The exercise results were stored in an SQL database light within the simulator app itself after each test.

SIMISGEST-VR Simulator Cost

The Leap Motion Controller costs approximately US $130, and the hardware elements cost approximately US $70. LapSim essence (Surgical Science Sweden AB) is a portable VR simulator that enables people to learn basic skills. It does not include haptics and is not available for sale, but it can be hired for 6 months at an approximate price of US $5500. To date,

there are no publications about the validity of the tasks that this device proposes.

Discussion

General

The aim of this study was to evaluate a simulation instrument—SIMISGEST-VR—and to document the sources of validity evidence for task scores, relations to other variables, internal structure, consequences of testing, and response process.

Technology-enhanced simulation is defined as an "educational tool or device with which the learner physically interacts to mimic an aspect of clinical care for the purpose of teaching or assessment [83,84]." The use of simulators for learning basic psychomotor skills in minimally invasive surgery has been supported by multiple systematic reviews [16,17,85-90].

In the current state-of-the-art conceptual framework, validity is defined as the appropriate interpretation or use of test results and therefore applies only to the scores or interpretation in a

specific context. The commonly used term *valid instrument* is inaccurate, and validity must be established for each intended interpretation [46,48,50]. Thus, when an evaluation instrument is said to be "valid" or to have been "validated," it is essential to take into account the learning context, the performance context, the domain content, and the exigency of decisions taken on the basis of test results [91].

Validation refers to the process of collecting validity evidence to evaluate the appropriateness of the interpretations, uses, and decisions based on assessment results [52]. Validation is, therefore, a process and not an endpoint, and it involves gathering evidence and taking decisions based on the interpretation of the data obtained. In our case, validation required a series of experiments designed to provide evidence that the scores measured in SIMISGEST-VR reflected the technical skills they purported to measure [92].

The first step in any validity evaluation is to clearly define the construct. The construct we focused on validating was training and learning basic psychomotor skills in minimally invasive surgery using a low-cost portable simulator called SIMISGEST-VR. Several systematic reviews [16,93-96] have found that basic psychomotor skills can be learned in low-cost simulation models; however, low-cost simulators are often box trainers made from cardboard boxes [97], plastic crates [98], folding portable boxes [99], and boxes that require the use of laparoscopic equipment [100,101] or even an iPad [102]. There are no low-cost VR simulators on the market.

An important finding from this study was the high percentage of surgical residents and surgeons that had no experience with simulators, and the very low percentage of surgical residents who had experience with hybrid and VR simulators. This finding can be explained by the high cost of this type of simulator, which, in many countries, prevents the creation of simulation centers for learning basic psychomotor skills in minimally invasive surgery and constitutes an argument in favor of exploring the development of models of low-cost portable VR simulators such as SIMISGEST-VR. Ucelli [44] demonstrated a comparable outcome between supervised simulator practice and unstructured free simulator access without mentoring and, therefore, that "take home" simulation was both viable and economically beneficial.

Validity Hypothesis 1: Relations to Other Variables

It is currently considered that a comparison between reference individuals and experts does not constitute an important validity argument [103,104]. However, it is the type of evidence for relations to other variables that is most often referred to in the literature [105]. The SIMISGEST-VR tasks were unable to demonstrate any difference in the performance scores between the reference group and the experts. A statistically significant difference was found between these 2 groups only in the time taken to complete tasks 3 (P=.006) and 6 (P=.02), which were the most complex.

Although some studies support the hypothesis that videogaming experience has a positive impact on minimally invasive surgery performance [106-113]. In this study, a significant difference was found for the reference group only in task 5 (diathermy; P=.003); in the other tasks, prior experience did not have any impact on performance. The demographic characterization made it clear that frequent videogaming (daily or weekly) was low in all population groups, which can explain the absence of impact on performance.

The lack of evidence for relations to other variables in this study can also be explained by the simplicity and ease of the proposed tasks.

Validity Hypothesis 2: Internal Structure

The items demonstrated high internal consistency (Cronbach α=.81). A test should not be used if it has a Cronbach α<0.7, and it should not be used for important decisions on an individual unless the Cronbach α>0.9 [57,114-116]. In our case, therefore, the result enables us to support the use of SIMISGEST-VR tasks as a self-assessment test used for formative assessment [54].

Test-retest reliability is the correlation between scores for a test administered more than once among a homogeneous group of test takers at 2 different times (temporal stability); the longer the period of time, the less likely it is that a person will remember the simulator tasks and, therefore, the greater the test-retest threat will be [116-119]. In this study, the second exercise was performed 6 months after the first one, and the results obtained demonstrate significant evidence for the temporal stability of scores in the 6 tasks. When the metric used was time, similar results were obtained in all but tasks 4 (P=.048) and 6 (P=.03) when comparing the tenth trial.

Validity Hypothesis 3: Evidence for Consequences of Testing

The most important finding of this study is that that the reference group learned in all the SIMISGEST-VR tasks. Excess diathermy error, defined as a contact time longer than 2 seconds from the moment of initial contact, fell significantly (P=.003) between the first and tenth trials for task 5 in the reference group, which also constitutes evidence for a learning curve. The experts group achieved a learning curve in all the tasks when time was taken as a metric, and for tasks 1 (P<.001), 3 (P=.03), and 4 (P=.01) when the final score of the test was taken into account. We, therefore, consider that the SIMISGEST-VR tasks can be used for the purpose of enabling novices without any prior experience to learn basic psychomotor skills in minimally invasive surgery.

This study has several strengths. The reference group sample included 100 students from 2 faculties of medicine, one public and one private; surgeons from a range of specialties; and surgical residents in general surgery and obstetric-gynecologic surgery. Physical simulators require the presence of a specialized tutor, a scarce, high-cost human resource, whereas VR simulators provide metrics and automatic feedback and allow the physical presence of a tutor to be dispensed with. At times of a pandemic such as COVID-19, this concept of education via VR takes on considerable significance because it avoids the need for learners to travel to simulation laboratories and, therefore, avoids close contact between students and instructors. This study also has limitations. Although the size of the reference group was large, a larger expert group would have

been desirable. The sample size in our study was one of availability; as such, there are relatively more participants with minimal surgical experience compared to those with a lot of experience, such as senior surgical residents and surgeons. The low number of senior residents prevented significant results from being obtained when comparing them to the other groups. Another limitation of the data analysis in this study is that there was no statistical analysis performed before the trial to evaluate the proper sample size or to determine the Likert scale.

Conclusions

This study has provided evidence to support the use of SIMISGEST-VR as a low-cost portable tool for the purpose of enabling novices without any prior experience to learn basic psychomotor skills in minimally invasive surgery. The tasks

for learning basic motor skills in minimally invasive surgery demonstrated high internal consistency and high test-retest reliability among the reference group when assessing the task scores. The expert group also managed to obtain a learning curve in all the tasks when assessing the time metric. In this study, we were able to demonstrate partial evidence for relations to other variables and strong evidence for internal structure and test consequences.

Future work streams include the creation of different levels of difficulty in the tasks. We also intend to develop an app that can be downloaded online, which contains the full training program. Finally, we hope to develop simulation models using the Leap Motion Controller and other gesture-recognition devices such as the Myo armband.

Authors' Contributions

All authors contributed substantially to the study conception and design, data analysis, interpretation of the findings, and manuscript drafting. FÁL participated in the collection and assembly of data. FA participated in the data analysis and statistical modeling. FS-R is the guarantor of the paper. All authors read, revised, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Form for the demographic questionnaire.
[DOCX File , 15 KB - games_v8i4e19723_app1.docx]

References

1. Litynski GS. Profiles in laparoscopy: Mouret, Dubois, and Perissat: the laparoscopic breakthrough in Europe (1987-1988). JSLS 1999;3(2):163-167 [FREE Full text] [Medline: 10444020]
2. Yamashita Y, Kurohiji T, Kakegawa T. Evaluation of two training programs for laparoscopic cholecystectomy: incidence of major complications. World J Surg 1994;18(2):279-85; discussion 285. [doi: 10.1007/BF00294415] [Medline: 8042335]
3. Archer SB, Brown DW, Smith CD, Branum GD, Hunter JG. Bile duct injury during laparoscopic cholecystectomy: results of a national survey. Ann Surg 2001 Oct;234(4):549-58; discussion 558. [doi: 10.1097/00000658-200110000-00014] [Medline: 11573048]
4. Berci G. Complications of laparoscopic cholecystectomy. Surg Endosc 1998 Apr;12(4):291-293. [doi: 10.1007/s004649900656] [Medline: 9543515]
5. Rogers DA, Elstein AS, Bordage G. Improving continuing medical education for surgical techniques: applying the lessons learned in the first decade of minimal access surgery. Ann Surg 2001 Feb;233(2):159-166. [doi: 10.1097/00000658-200102000-00003] [Medline: 11176120]
6. Hugh TB. New strategies to prevent laparoscopic bile duct injury--surgeons can learn from pilots. Surgery 2002 Nov;132(5):826-835. [doi: 10.1067/msy.2002.127681] [Medline: 12464867]
7. MacFadyen BV, Vecchio R, Ricardo AE, Mathis CR. Bile duct injury after laparoscopic cholecystectomy. the United States experience. Surg Endosc 1998 Apr 10;12(4):315-321. [doi: 10.1007/s004649900661] [Medline: 9543520]
8. Wolfe BM, Gardiner B, Frey CF. Laparoscopic cholecystectomy: a remarkable development. JAMA 2015 Oct 06;314(13):1406. [doi: 10.1001/jama.2014.12014] [Medline: 26441196]
9. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 2002 Oct;236(4):458-63; discussion 463. [doi: 10.1097/00000658-200210000-00008] [Medline: 12368674]
10. Gallagher AG, Richie K, McClure N, McGuigan J. Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality. World J Surg 2001 Nov;25(11):1478-1483. [doi: 10.1007/s00268-001-0133-1] [Medline: 11760752]
11. Gallagher AG, Smith CD, Bowers SP, Seymour NE, Pearson A, McNatt S, et al. Psychomotor skills assessment in practicing surgeons experienced in performing advanced laparoscopic procedures. J Am Coll Surg 2003 Sep;197(3):479-488. [doi: 10.1016/S1072-7515(03)00535-0] [Medline: 12946803]

XSL•FO

RenderX

12. Korndorffer JR, Dunne JB, Sierra R, Stefanidis D, Touchard CL, Scott DJ. Simulator training for laparoscopic suturing using performance goals translates to the operating room. J Am Coll Surg 2005 Jul;201(1):23-29. [doi: 10.1016/j.jamcollsurg.2005.02.021] [Medline: 15978440]

13. Aggarwal R, Ward J, Balasundaram I, Sains P, Athanasiou T, Darzi A. Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. Ann Surg 2007 Nov;246(5):771-779. [doi: 10.1097/SLA.0b013e3180f61b09] [Medline: 17968168]

14. Sturm LP, Windsor JA, Cosman PH, Cregan P, Hewett PJ, Maddern GJ. A systematic review of skills transfer after surgical simulation training. Ann Surg 2008 Aug;248(2):166-179. [doi: 10.1097/SLA.0b013e318176bf24] [Medline: 18650625]

15. Seymour NE. VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. World J Surg 2008 Feb 3;32(2):182-188. [doi: 10.1007/s00268-007-9307-9] [Medline: 18060453]

16. Zendejas B, Brydges R, Hamstra SJ, Cook DA. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. Ann Surg 2013 Apr;257(4):586-593. [doi: 10.1097/SLA.0b013e318288c40b] [Medline: 23407298]

17. Dawe SR, Pena GN, Windsor JA, Broeders JAJL, Cregan PC, Hewett PJ, et al. Systematic review of skills transfer after surgical simulation-based training. Br J Surg 2014 Aug;101(9):1063-1076. [doi: 10.1002/bjs.9482] [Medline: 24827930]

18. Hyltander A, Liljegren E, Rhodin P, Lönroth H. The transfer of basic skills learned in a laparoscopic simulator to the operating room. Surg Endosc 2002 Sep;16(9):1324-1328. [doi: 10.1007/s00464-001-9184-5] [Medline: 11988802]

19. Youngblood PL, Srivastava S, Curet M, Heinrichs WL, Dev P, Wren SM. Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance. J Am Coll Surg 2005 Apr;200(4):546-551. [doi: 10.1016/j.jamcollsurg.2004.11.011] [Medline: 15804468]

20. Sutton C, McCloy R, Middlebrook A, Chater P, Wilson M, Stone R. MIST VR. A laparoscopic surgery procedures trainer and evaluator. Stud Health Technol Inform 1997;39:598-607. [Medline: 10173070]

21. Ahlberg G, Heikkinen T, Iselius L, Leijonmarck C, Rutqvist J, Arvidsson D. Does training in a virtual reality simulator improve surgical performance? Surg Endosc 2002 Jan 12;16(1):126-129. [doi: 10.1007/s00464-001-9025-6] [Medline: 11961622]

22. Debes AJ, Aggarwal R, Balasundaram I, Jacobsen MB. A tale of two trainers: virtual reality versus a video trainer for acquisition of basic laparoscopic skills. Am J Surg 2010 Jun;199(6):840-845. [doi: 10.1016/j.amjsurg.2009.05.016] [Medline: 20079480]

23. Kothari SN, Kaplan BJ, DeMaria EJ, Broderick TJ, Merrell RC. Training in laparoscopic suturing skills using a new computer-based virtual reality simulator (MIST-VR) provides results comparable to those with an established pelvic trainer system. J Laparoendosc Adv Surg Tech A 2002 Jun;12(3):167-173. [doi: 10.1089/10926420260188056] [Medline: 12184901]

24. Torkington J, Smith S, Rees B, Darzi A. Skill transfer from virtual reality to a real laparoscopic task. Surg Endosc 2001 Oct 15;15(10):1076-1079. [doi: 10.1007/s004640000233] [Medline: 11727073]

25. Torkington J, Smith S, Rees B, Darzi A. The role of the basic surgical skills course in the acquisition and retention of laparoscopic skill. Surg Endosc 2001 Oct 15;15(10):1071-1075. [doi: 10.1007/s004640000183] [Medline: 11727072]

26. Alvarez-Lopez F, Maina MF, Saigí-Rubió F. Use of commercial off-the-shelf devices for the detection of manual gestures in surgery: systematic literature review. J Med Internet Res 2019 Apr 14;21(5):e11925 [FREE Full text] [doi: 10.2196/11925] [Medline: 31066679]

27. Ogura T, Sato M, Ishida Y, Hayashi N, Doi K. Development of a novel method for manipulation of angiographic images by use of a motion sensor in operating rooms. Radiol Phys Technol 2014 Jul;7(2):228-234. [doi: 10.1007/s12194-014-0259-0] [Medline: 24609904]

28. Mauser S, Burgert O. Touch-free, gesture-based control of medical devices and software based on the leap motion controller. Stud Health Technol Inform 2014;196:265-270. [Medline: 24732520]

29. Bachmann D, Weichert F, Rinkenauer G. Evaluation of the leap motion controller as a new contact-free pointing device. Sensors (Basel) 2014 Dec 24;15(1):214-233 [FREE Full text] [doi: 10.3390/s150100214] [Medline: 25609043]

30. Weichert F, Bachmann D, Rudak B, Fisseler D. Analysis of the accuracy and robustness of the leap motion controller. Sensors (Basel) 2013 May 14;13(5):6380-6393 [FREE Full text] [doi: 10.3390/s130506380] [Medline: 23673678]

31. Guna J, Jakus G, Pogačnik M, Tomažič S, Sodnik J. An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. Sensors (Basel) 2014 Feb;14(2):3702-3720 [FREE Full text] [doi: 10.3390/s140203702] [Medline: 24566635]

32. Alvarez-Lopez F, Maina MF, Saigí-Rubió F. Natural User Interfaces: Is It a Solution to Accomplish Ubiquitous Training in Minimally Invasive Surgery? Surg Innov 2016 Aug 09;23(4):429-430. [doi: 10.1177/1553350616639145] [Medline: 27009688]

33. Oropesa I, de Jong T, Sánchez-González P, Dankelman J, Gómez E. Feasibility of tracking laparoscopic instruments in a box trainer using a Leap Motion Controller. Measurement 2016 Feb;80:115-124. [doi: 10.1016/j.measurement.2015.11.018]

34. Beck P. Accurate three-dimensional instrument positioning. Free Patents Online. 2016. URL: http://www.freepatentsonline.com/20160354152.pdf [accessed 2020-10-15]

35. Lahanas V, Loukas C, Georgiou K, Lababidi H, Al-Jaroudi D. Virtual reality-based assessment of basic laparoscopic skills using the Leap Motion controller. Surg Endosc 2017 Dec 2;31(12):5012-5023. [doi: 10.1007/s00464-017-5503-3] [Medline: 28466361]

36.   Travaglini T, Swaney P, Weaver K, Webster R. Initial experiments with the leap motion as a user interface in robotic endonasal surgery. Robot Mechatron (2015) 2016;37:171-179 [FREE Full text] [doi: 10.1007/978-3-319-22368-1_17] [Medline: 26752501]

37.   Juanes JA, Gómez JJ, Peguero PD, Ruisoto P. Digital environment for movement control in surgical skill training. J Med Syst 2016 Jun 18;40(6):133. [doi: 10.1007/s10916-016-0495-4] [Medline: 27091754]

38.   Partridge RW, Brown FS, Brennan PM, Hennessey IAM, Hughes MA. The LEAPTM gesture interface device and take-home laparoscopic simulators: a study of construct and concurrent validity. Surg Innov 2016 Feb 14;23(1):70-77. [doi: 10.1177/1553350615594734] [Medline: 26178693]

39.   Wright T, de Ribaupierre S, Eagleson R. Design and evaluation of an augmented reality simulator using leap motion. Healthc Technol Lett 2017 Oct;4(5):210-215 [FREE Full text] [doi: 10.1049/htl.2017.0070] [Medline: 29184667]

40.   Botden SMBI, Jakimowicz JJ. What is going on in augmented reality simulation in laparoscopic surgery? Surg Endosc 2009 Aug;23(8):1693-1700 [FREE Full text] [doi: 10.1007/s00464-008-0144-1] [Medline: 18813987]

41.   Papanikolaou IG. Assessment of medical simulators as a training programme for current surgical education. Hellenic J Surg 2013 Aug 9;85(4):240-248. [doi: 10.1007/s13126-013-0047-z]

42.   Hennessey IA, Hewett P. Construct, concurrent, and content validity of the eoSim laparoscopic simulator. J Laparoendosc Adv Surg Tech A 2013 Oct;23(10):855-860. [doi: 10.1089/lap.2013.0229] [Medline: 23968255]

43.   Hruby GW, Sprenkle PC, Abdelshehid C, Clayman RV, McDougall EM, Landman J. The EZ Trainer: validation of a portable and inexpensive simulator for training basic laparoscopic skills. J Urol 2008 Feb;179(2):662-666. [doi: 10.1016/j.juro.2007.09.030] [Medline: 18082210]

44.   Uccelli J, Kahol K, Ashby A, Smith M, Ferrara J. The validity of take-home surgical simulators to enhance resident technical skill proficiency. Am J Surg 2011 Mar;201(3):315-9; discussion 319. [doi: 10.1016/j.amjsurg.2010.08.028] [Medline: 21367370]

45.   Estándares Para Pruebas Educativas y Psicológicas. JSTOR.: American Educational Research Association; 2018. URL: www.jstor.org/stable/j.ctvr43hg2 [accessed 2020-10-17]

46.   American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington, DC: American Educational Research Association; 2014.

47.   Downing SM. Validity: on meaningful interpretation of assessment data. Med Educ 2003 Sep;37(9):830-837. [doi: 10.1046/j.1365-2923.2003.01594.x] [Medline: 14506816]

48.   Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med 2006 Feb;119(2):166.e7-166.16. [doi: 10.1016/j.amjmed.2005.10.036] [Medline: 16443422]

49.   Cook DA, Zendejas B, Hamstra SJ, Hatala R, Brydges R. What counts as validity evidence? Examples and prevalence in a systematic review of simulation-based assessment. Adv Health Sci Educ Theory Pract 2014 May;19(2):233-250. [doi: 10.1007/s10459-013-9458-4] [Medline: 23636643]

50.   Messick S. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist 1995 Sep;50(9):741-749. [doi: 10.1037/0003-066x.50.9.741]

51.   Messick S. Meaning and values in test validation: the science and ethics of assessment. Educational Researcher 2016 Jul;18(2):5-11. [doi: 10.3102/0013189x018002005]

52.   Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. Adv Simul (Lond) 2016;1:31 [FREE Full text] [doi: 10.1186/s41077-016-0033-y] [Medline: 29450000]

53.   Alvarez-Lopez F, Maina MF, Saigí-Rubió F. Use of a low-cost portable 3d virtual reality gesture-mediated simulator for training and learning basic psychomotor skills in minimally invasive surgery: development and content validity study. J Med Internet Res 2020 Jul 14;22(7):e17491 [FREE Full text] [doi: 10.2196/17491] [Medline: 32673217]

54.   Ghaderi I, Manji F, Park YS, Juul D, Ott M, Harris I, et al. Technical skills assessment toolbox: a review using the unitary framework of validity. Ann Surg 2015 Feb;261(2):251-262. [doi: 10.1097/SLA.0000000000000520] [Medline: 24424150]

55.   Beckman TJ, Cook DA, Mandrekar JN. What is the validity evidence for assessments of clinical teaching? J Gen Intern Med 2005 Dec;20(12):1159-1164. [doi: 10.1111/j.1525-1497.2005.0258.x] [Medline: 16423109]

56.   Sweet R, Hananel D, Lawrenz F. A unified approach to validation, reliability, and education study design for surgical technical skills training. Arch Surg 2010 Feb;145(2):197-201. [doi: 10.1001/archsurg.2009.266] [Medline: 20157089]

57.   Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. Surg Endosc 2003 Oct 1;17(10):1525-1529. [doi: 10.1007/s00464-003-0035-4] [Medline: 14502403]

58.   Worster A, Haines T. Advanced statistics: understanding medical record review (MRR) studies. Acad Emerg Med 2004 Feb;11(2):187-192. [Medline: 14759964]

59.   Warner B. The sciences of the artificial. Journal of the Operational Research Society 2017 Dec 19;20(4):509-510 [FREE Full text] [doi: 10.1057/jors.1969.121]

60.   Manson N. Is operations research really research? ORiON 2006 Dec 01;22(2). [doi: 10.5784/22-2-40]

61.   Dresch A, Pacheco-Lacerda D, Valle-Antunes J. Design science research. In: A Method for Science and Technology Advancement. Switzerland: Springer International Publishing; 2015.

XSL·FO

RenderX

62. Lacerda DP, Dresch A, Proença A, Antunes Júnior JAV. Design Science Research: método de pesquisa para a engenharia de produção. Gest. Prod 2013 Nov 26;20(4):741-761. [doi: 10.1590/s0104-530x2013005000014]

63. Hevner A, Chatterjee S. Design research in information systems. In: Theory and Practice. Boston, MA: Springer US; 2010.

64. Wilson M, Middlebrook A, Sutton C, Stone R, McCloy R. MIST VR: a virtual reality trainer for laparoscopic surgery assesses performance. Ann R Coll Surg Engl 1997 Nov;79(6):403-404 [FREE Full text] [Medline: 9422863]

65. Ali M, Mowery Y, Kaplan B, DeMaria E. Training the novice in laparoscopy. More challenge is better. Surg Endosc 2002 Dec 1;16(12):1732-1736. [doi: 10.1007/s00464-002-8850-6] [Medline: 12140638]

66. Grantcharov TP, Rosenberg J, Pahle E, Funch-Jensen P. Virtual reality computer simulation. Surg Endosc 2001 Mar;15(3):242-244. [doi: 10.1007/s004640090008] [Medline: 11344422]

67. Hamilton E, Scott D, Fleming J, Rege R, Laycock R, Bergen P, et al. Comparison of video trainer and virtual reality training systems on acquisition of laparoscopic skills. Surg Endosc 2002 Mar;16(3):406-411. [doi: 10.1007/s00464-001-8149-z] [Medline: 11928017]

68. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Br J Surg 2004 Feb 29;91(2):146-150. [doi: 10.1002/bjs.4407] [Medline: 14760660]

69. Gonzalez R, Bowers SP, Smith CD, Ramshaw BJ. Does setting specific goals and providing feedback during training result in better acquisition of laparoscopic skills? Am Surg 2004 Jan;70(1):35-39. [Medline: 14964544]

70. McClusky DA, Gallagher AG, Ritter E, Lederman AB, Van Sickle KR, Baghai M, et al. Virtual reality training improves junior residents' operating room performance: Results of a prospective, randomized, double-blinded study of the complete laparoscopic cholecystectomy. Journal of the American College of Surgeons 2004 Sep;199(3):73. [doi: 10.1016/j.jamcollsurg.2004.05.157]

71. Grantcharov TP, Funch-Jensen P. Can everyone achieve proficiency with the laparoscopic technique? Learning curve patterns in technical skills acquisition. Am J Surg 2009 Apr;197(4):447-449. [doi: 10.1016/j.amjsurg.2008.01.024] [Medline: 19217604]

72. Gallagher AG, Seymour NE, Jordan-Black J, Bunting BP, McGlade K, Satava RM. Prospective, Randomized Assessment of Transfer of Training (ToT) and Transfer Effectiveness Ratio (TER) of Virtual Reality Simulation Training for Laparoscopic Skill Acquisition. Annals of Surgery 2013;257(6):1025-1031. [doi: 10.1097/sla.0b013e318284f658]

73. Badash I, Burtt K, Solorzano CA, Carey JN. Innovations in surgery simulation: a review of past, current and future techniques. Ann Transl Med 2016 Dec;4(23):453-453 [FREE Full text] [doi: 10.21037/atm.2016.12.24] [Medline: 28090509]

74. Taffinder N, Sutton C, Fishwick RJ, McManus IC, Darzi A. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator. Stud Health Technol Inform 1998;50:124-130. [Medline: 10180527]

75. Chaudhry A, Sutton C, Wood J, Stone R, McCloy R. Learning rate for laparoscopic surgical skills on MIST VR, a virtual reality simulator: quality of human-computer interface. Ann R Coll Surg Engl 1999 Jul;81(4):281-286 [FREE Full text] [Medline: 10615201]

76. Stylopoulos N, Vosburgh KG. Assessing technical skill in surgery and endoscopy: a set of metrics and an algorithm (C-PASS) to assess skills in surgical and endoscopic procedures. Surg Innov 2007 Jun;14(2):113-121. [doi: 10.1177/1553350607302330] [Medline: 17558017]

77. Hatala R, Cook DA, Zendejas B, Hamstra SJ, Brydges R. Feedback for simulation-based procedural skills training: a meta-analysis and critical narrative synthesis. Adv Health Sci Educ Theory Pract 2014 May;19(2):251-272. [doi: 10.1007/s10459-013-9462-8] [Medline: 23712700]

78. Gallagher A, Satava R. Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. Learning curves and reliability measures. Surg Endosc 2002 Dec 1;16(12):1746-1752. [doi: 10.1007/s00464-001-8215-6] [Medline: 12140641]

79. Archer J. State of the science in health professional education: effective feedback. Med Educ 2010 Jan;44(1):101-108. [doi: 10.1111/j.1365-2923.2009.03546.x] [Medline: 20078761]

80. Walsh CM, Ling SC, Wang CS, Carnahan H. Concurrent versus terminal feedback: it may be better to wait. Academic Medicine 2009;84(Supplement):S54-S57. [doi: 10.1097/acm.0b013e3181b38daf] [Medline: 19907387]

81. Grantcharov TP, Schulze S, Kristiansen VB. The impact of objective assessment and constructive feedback on improvement of laparoscopic performance in the operating room. Surg Endosc 2007 Dec 13;21(12):2240-2243. [doi: 10.1007/s00464-007-9356-z] [Medline: 17440781]

82. Stefanidis D, Korndorffer JR, Heniford BT, Scott DJ. Limited feedback and video tutorials optimize learning and resource utilization during laparoscopic simulator training. Surgery 2007 Aug;142(2):202-206. [doi: 10.1016/j.surg.2007.03.009] [Medline: 17689686]

83. Cook DA, Brydges R, Zendejas B, Hamstra SJ, Hatala R. Technology-enhanced simulation to assess health professionals: a systematic review of validity evidence, research methods, and reporting quality. Acad Med 2013 Jun;88(6):872-883. [doi: 10.1097/ACM.0b013e31828ffdcf] [Medline: 23619073]

84. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 07;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]

85. Haque S, Srinivasan S. A meta-analysis of the training effectiveness of virtual reality surgical simulators. IEEE Trans Inf Technol Biomed 2006 Jan;10(1):51-58. [doi: 10.1109/titb.2005.855529] [Medline: 16445249]

86. Brydges R, Hatala R, Zendejas B, Erwin PJ, Cook DA. Linking simulation-based educational assessments and patient-related outcomes: a systematic review and meta-analysis. Acad Med 2015 Feb;90(2):246-256. [doi: 10.1097/ACM.0000000000000549] [Medline: 25374041]

87. Nagendran M, Gurusamy KS, Aggarwal R, Loizidou M, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. Cochrane Database Syst Rev 2013 Aug 27(8):CD006575 [FREE Full text] [doi: 10.1002/14651858.CD006575.pub3] [Medline: 23980026]

88. Thomas MP. The role of simulation in the development of technical competence during surgical training: a literature review. Int J Med Educ 2013 Mar 16;4:48-58. [doi: 10.5116/ijme.513b.2df7]

89. Dawe SR, Windsor JA, Broeders JA, Cregan PC, Hewett PJ, Maddern GJ. A systematic review of surgical skills transfer after simulation-based training: laparoscopic cholecystectomy and endoscopy. Ann Surg 2014 Feb;259(2):236-248. [doi: 10.1097/SLA.0000000000000245] [Medline: 24100339]

90. Gurusamy K, Aggarwal R, Palanivelu L, Davidson BR. Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. Br J Surg 2008 Sep;95(9):1088-1097. [doi: 10.1002/bjs.6344] [Medline: 18690637]

91. Andreatta P, Gruppen L. Conceptualising and classifying validity evidence for simulation. Med Educ 2009 Nov;43(11):1028-1035. [doi: 10.1111/j.1365-2923.2009.03454.x] [Medline: 19874494]

92. Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, et al. Proving the value of simulation in laparoscopic surgery. Ann Surg 2004 Sep;240(3):518-25; discussion 525. [doi: 10.1097/01.sla.0000136941.46529.56] [Medline: 15319723]

93. Li MM, George J. A systematic review of low-cost laparoscopic simulators. Surg Endosc 2017 Jan;31(1):38-48 [FREE Full text] [doi: 10.1007/s00464-016-4953-3] [Medline: 27194266]

94. Nguyen T, Braga LH, Hoogenes J, Matsumoto ED. Commercial video laparoscopic trainers versus less expensive, simple laparoscopic trainers: a systematic review and meta-analysis. J Urol 2013 Sep;190(3):894-899. [doi: 10.1016/j.juro.2013.03.115] [Medline: 23567747]

95. Nagendran M, Toon C, Davidson B, Gurusamy K. Laparoscopic surgical box model training for surgical trainees with no prior laparoscopic experience. Cochrane Database Syst Rev 2014 Jan 17(1):CD010479. [doi: 10.1002/14651858.CD010479.pub2] [Medline: 24442763]

96. Diesen DL, Erhunmwunsee L, Bennett KM, Ben-David K, Yurcisin B, Ceppa EP, et al. Effectiveness of laparoscopic computer simulator versus usage of box trainer for endoscopic surgery training of novices. J Surg Educ 2011 Jul;68(4):282-289. [doi: 10.1016/j.jsurg.2011.02.007] [Medline: 21708364]

97. Blacker AJ. How to build your own laparoscopic trainer. J Endourol 2005 Jul;19(6):748-752. [doi: 10.1089/end.2005.19.748] [Medline: 16053370]

98. Smith MD, Norris JM, Kishikova L, Smith DP. Laparoscopic simulation for all: two affordable, upgradable, and easy-to-build laparoscopic trainers. J Surg Educ 2013 Mar;70(2):217-223. [doi: 10.1016/j.jsurg.2012.11.005] [Medline: 23427967]

99. Nakamura LY, Martin GL, Fox JC, Andrews PE, Humphreys M, Castle EP. Comparing the portable laparoscopic trainer with a standardized trainer in surgically naïve subjects. J Endourol 2012 Jan;26(1):67-72. [doi: 10.1089/end.2011.0335] [Medline: 21999424]

100. Ricchiuti D, Ralat DA, Evancho-Chapman M, Wyneski H, Cerone J, Wegryn JD. A simple cost-effective design for construction of a laparoscopic trainer. J Endourol 2005 Oct;19(8):1000-2; discussion 1002. [doi: 10.1089/end.2005.19.1000] [Medline: 16253069]

101. Mughal M. A cheap laparoscopic surgery trainer. Ann R Coll Surg Engl 1992 Jul;74(4):256-257 [FREE Full text] [Medline: 1416677]

102. Ruparel RK, Brahmbhatt RD, Dove JC, Hutchinson RC, Stauffer JA, Bowers SP, et al. "iTrainers"--novel and inexpensive alternatives to traditional laparoscopic box trainers. Urology 2014 Jan;83(1):116-120. [doi: 10.1016/j.urology.2013.09.030] [Medline: 24246314]

103. Cook DA. Much ado about differences: why expert-novice comparisons add little to the validity argument. Adv Health Sci Educ Theory Pract 2015 Aug;20(3):829-834. [doi: 10.1007/s10459-014-9551-3] [Medline: 25260974]

104. Zendejas B, Ruparel RK, Cook DA. Validity evidence for the Fundamentals of Laparoscopic Surgery (FLS) program as an assessment tool: a systematic review. Surg Endosc 2016 Feb;30(2):512-520. [doi: 10.1007/s00464-015-4233-7] [Medline: 26091982]

105. Borgersen NJ, Naur TMH, Sørensen SMD, Bjerrum F, Konge L, Subhi Y, et al. Gathering validity evidence for surgical simulation: a systematic review. Ann Surg 2018 Jun;267(6):1063-1068. [doi: 10.1097/SLA.0000000000002652] [Medline: 29303808]

106. Rosser JC, Lynch PJ, Cuddihy L, Gentile DA, Klonsky J, Merrell R. The impact of video games on training surgeons in the 21st century. Arch Surg 2007 Feb 01;142(2):181-6; discusssion 186. [doi: 10.1001/archsurg.142.2.181] [Medline: 17309970]

107. van Dongen KW, Verleisdonk EMM, Schijven MP, Broeders IAMJ. Will the Playstation generation become better endoscopic surgeons? Surg Endosc 2011 Jul 17;25(7):2275-2280 [FREE Full text] [doi: 10.1007/s00464-010-1548-2] [Medline: 21416186]

108. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Impact of hand dominance, gender, and experience with computer games on performance in virtual reality laparoscopy. Surg Endosc 2003 Jul 1;17(7):1082-1085. [doi: 10.1007/s00464-002-9176-0] [Medline: 12728373]

109. Boyle E, Kennedy A, Traynor O, Hill ADK. Training surgical skills using nonsurgical tasks--can Nintendo Wii™ improve surgical performance? J Surg Educ 2011;68(2):148-154. [doi: 10.1016/j.jsurg.2010.11.005] [Medline: 21338974]

110. Harper JD, Kaiser S, Ebrahimi K, Lamberton GR, Hadley HR, Ruckle HC, et al. Prior video game exposure does not enhance robotic surgical performance. J Endourol 2007 Oct;21(10):1207-1210. [doi: 10.1089/end.2007.9905] [Medline: 17949327]

111. Glaser AY, Hall CB, Uribe JI, Fried MP. The effects of previously acquired skills on sinus surgery simulator performance. Otolaryngol Head Neck Surg 2005 Oct;133(4):525-530. [doi: 10.1016/j.otohns.2005.06.022] [Medline: 16213923]

112. Madan AK, Frantzides CT, Park WC, Tebbit CL, Kumari NVA, O'Leary PJ. Predicting baseline laparoscopic surgery skills. Surg Endosc 2005 Jan;19(1):101-104. [doi: 10.1007/s00464-004-8123-7] [Medline: 15531975]

113. Kennedy A, Boyle E, Traynor O, Walsh T, Hill A. Video gaming enhances psychomotor skills but not visuospatial and perceptual abilities in surgical trainees. J Surg Educ 2011 Sep;68(5):414-420. [doi: 10.1016/j.jsurg.2011.03.009] [Medline: 21821223]

114. Wanzel KR, Ward M, Reznick RK. Teaching the surgical craft: from selection to certification. Curr Probl Surg 2002 Jun;39(6):573-659. [doi: 10.1067/mog.2002.123481] [Medline: 12037512]

115. Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychological Bulletin 1955;52(4):281-302. [doi: 10.1037/h0040957]

116. Feldman LS, Sherman V, Fried GM. Using simulators to assess laparoscopic competence: ready for widespread use? Surgery 2004 Jan;135(1):28-42. [doi: 10.1016/s0039-6060(03)00155-7] [Medline: 14694298]

117. Straub D, Gefen D. Validation Guidelines for IS Positivist Research. CAIS 2004;13(1):427. [doi: 10.17705/1CAIS.01324]

118. McDougall EM. Validation of surgical simulators. J Endourol 2007 Mar;21(3):244-247. [doi: 10.1089/end.2007.9985] [Medline: 17444766]

119. Kowalewski K, Hendrie JD, Schmidt MW, Garrow CR, Bruckner T, Proctor T, et al. Development and validation of a sensor- and expert model-based training system for laparoscopic surgery: the iSurgeon. Surg Endosc 2017 May;31(5):2155-2165. [doi: 10.1007/s00464-016-5213-2] [Medline: 27604368]

Abbreviations

MIST-VR: Minimally Invasive Surgery Training - Virtual Reality
SIMISGEST-VR: Simulator of Minimally Invasive Surgery mediated by Gestures - Virtual Reality
VR: virtual reality

XSL•FO
RenderX

<u>Viewpoint</u>

The Impact of Artificial Intelligence on the Chess World

Delia Monica Duca Iliescu[1], BSc, MSc

Transilvania University of Brasov, Brasov, Romania

Corresponding Author:

Delia Monica Duca Iliescu, BSc, MSc

Transilvania University of Brasov

Bdul Eroilor 29

Brasov

Romania

Phone: 40 268413000

Email: delia.duca@unitbv.ro

Abstract

This paper focuses on key areas in which artificial intelligence has affected the chess world, including cheat detection methods, which are especially necessary recently, as there has been an unexpected rise in the popularity of online chess. Many major chess events that were to take place in 2020 have been canceled, but the global popularity of chess has in fact grown in recent months due to easier conversion of the game from offline to online formats compared with other games. Still, though a game of chess can be easily played online, there are some concerns about the increased chances of cheating. Artificial intelligence can address these concerns.

Introduction

All major chess events that were to take place in the second half of 2020 have been canceled, including the 44th Chess Olympiad and the match for the title of World Chess Champion. However, a lot of major national and international chess competitions were recently played online for the first time in history. Through the course of the COVID-19 pandemic, chess has been one of the few sports that has managed to remain relevant to the public and even gain popularity [1], with fans watching online chess tournaments from their homes.

Even though a game of chess can be played online in almost the same way it is played offline, there are some concerns about the increased possibilities of cheating, and many software companies are looking for a solution.

Context: Artificial Intelligence Progress Displayed in Chess

Artificial intelligence is undergoing a revolution, but at the same time, it has caused revolutionary changes in the world. Chess has inspired artificial intelligence progress for decades [2]. The developments in artificial intelligence for chess have advanced beyond gaming, changing the way machines and humans coexist. The progress of artificial intelligence, highlighted by

strategic games, has affected many other areas of interest, as has already been seen in recent years.

One of the most famous human versus machine events was the 1997 victory of Deep Blue, an IBM chess software, against the famous chess champion Garry Kasparov [3]. However, this victory was mainly achieved by brute force, since the Deep Blue program searched millions of positions per second to play chess at a slightly higher strength (the program won the 6-game match with Garry Kasparov with a difference of only 1 point), so it was probably more impressive that Kasparov, using human intuition, could still be almost as strong as a computer searching millions of positions per second. In the following years, computing power advanced to the point where even the best chess players had no chance of defeating a modern chess engine, as has been previously stated [4].

Significant progress in artificial intelligence–related fields was made only 20 years later when AlphaZero won a chess match in 2017 against Stockfish, one of the most powerful chess engines ever created, after a self-learning process of only 4 hours.

The AlphaZero algorithm did not try to use the brute force of computing power to identify as many moves as possible on the chessboard. Instead, using reinforcement learning, humans' learning process was mimicked by studying an impressive

XSL•FO

RenderX

number of chess games [5]. Such interaction with the environment is how people learn. Reinforcement learning, whereby an agent tries to maximize the reward in a "complex, uncertain environment" [6], is just a computational approach to interactive learning. The creators of AlphaZero claimed that the algorithm can learn to optimize decisions in any scenario without changes or guidance [7], and this was truly a breakthrough.

Furthermore, an algorithm that is even beyond AlphaZero was released just a year ago. On November 19, 2019, DeepMind launched the latest algorithm based on reinforcement learning, MuZero. The MuZero algorithm learned to play chess better than AlphaZero without even initially being told the rules of the game. The MuZero learning algorithm has a maximum of 1 million chess games saved in the buffer, with 3000 games played in parallel, as has been stated by Schrittwieser et al [8]. This was possible because DeepMind had access to Google's vast cloud infrastructure and used thousands of tensor processing unit chips specifically designed for neural network calculations.

There is no way to calculate every possible move on the chessboard up to the end of the game due to the number of possible positions in chess. The first tablebases that tried to solve at least the 5- to 6-piece chess endgames were developed by Ken Thompson [9]. The large amount of storage required to do this makes it impossible to realize this idea for all possible chess positions even to this day.

Thus, the AlphaZero program's learning of the game of chess in 4 hours with only the rules of the game was remarkable. Still, in the real world, these rules are rarely known. Therefore, the new MuZero algorithm, which manages to play chess on the same level as AlphaZero without receiving these rules, is even more spectacular. Later, even open-source engines such as Leela Chess Zero managed to reach the MuZero level and even surpass it. This was possible using computer power and improvements to the source code provided by a large number of volunteers around the world [10].

Artificial Intelligence Used for Cheat Detection in Chess Tournaments

It was thought for a long time that people would use artificial intelligence to learn how to play chess better, but now it is successfully used to detect if some contestants play better than they should, considering their game history. Still, these cheat detection mechanisms come with several controversial issues of their own.

As mentioned in the Anti-Cheating Guidelines published by the International Chess Federation (FIDE) in 2014, in most cases, a handheld metal detector is enough to ensure that electronic devices are not being carried into the playing venue, providing cheating protection for onboard games. However, for online chess games, cheat detection has proven to be much more difficult. Almost 4000 players coming from 55 European federations registered for the European Online Chess Championship that took place in May 2020, which was a record number of participants in an official international chess championship. A total of 5 out of 6 players at the top of the B

group (rating of 1400-1700) of the European Online Championship were disqualified. In total, more than 80 participants have been disqualified in all categories, which is approximately 2% of the players, the majority being from the beginners or youth categories. Numerous disqualifications at the first-ever European Online Chess Championship highlighted the greatest challenge faced by online chess: cheating [11].

Recently, various software companies, including DeepMind, which developed MuZero, have been working hard to improve existing cheat detection software or even to develop completely new software that will be able to estimate with almost 100% accuracy whether a player is cheating. For example, on Chess.com, a cheat detection system uses millions of chess games stored in its database to create a statistical model that assesses the low probability that a human player will match the top choices of an engine or even surpass the games of some of the greatest chess players in history. All reports of possible cheating are then carefully analyzed by a team of experts. The results are published every month in the website's "Month in Review."

Most of the games played during this period have been rapid chess, and chess games played with a longer time per player will be an even bigger problem, as there will be more time for possible forbidden electronic assistance. FIDE has already approved a complex cheat detection technology and an artificial intelligence behavior-tracking module for the FIDE Online Arena games. These online competitions were previously considered separate from the onboard ones, but in the recent pandemic context, the distinction between offline and online has been diminished, with many official competitions being played online for the first time in history.

Cheat Detection Issues

There are some sensitive issues that should be considered regarding online playing and fraud detection. First, there is a possibility of a noncheating player being classified as a cheater. This would clearly cause serious problems in the career of such a chess player. It is difficult to make this type of analysis for a chess player in their first game ever recorded, but after playing several games in an official online tournament, these programs can easily identify the unlikely differences between normal plays and the plays in which chess engines are used.

Trying to use chess engines every time to play perfectly is not plausible either; this is also easy to detect, as the player would rank first, and in this case, the play can be evaluated by a human expert. Consequently, for those whose playing fluctuates because they sometimes rely on their own decisions and sometimes use chess engines, the new type of cheat detection software based on artificial intelligence detects this very well, with much higher speed and accuracy than a human referee. All suspected cases of fraud are then checked individually by a team of experts in the field before being made public.

The software uses a particular kind of algorithm to detect an unlikely chess move from a specific player. To catch an alleged cheater, the software assesses a set of chess positions played by a certain player (preferably at least a few hundred but the

XSL•FO

RenderX

analysis can also work with just a few games) and translates them into a personal performance rating. This helps determine a nonplausible move by using a machine learning algorithm based on clustering, and the result is a list of flags indicating cheating. Of course, when there is a higher number of chess games played by a specific player in the training set, the program will have higher accuracy.

Cheat detection algorithms must be applied carefully to top players because whenever one of them makes a novel move, it could trigger the fraud signal. Still, we can assume that the best players in the world would not risk their reputation over a tournament. It is thought that practicing chess develops certain intrinsic qualities related to fairness, which is probably why most performance chess players never cheat. However, the growing popularity of online chess, especially in recent months, has led to an increase in the number of people completely outside of the chess world trying to win major online tournaments, and some of these people have tried to use chess engines to play better than they normally would at the chessboard.

Conclusions

What is important for progress in the field of artificial intelligence is that we can measure any breakthrough in machine learning algorithms, for example, through the chess engines that are using them. In other words, it could be argued that chess is a battleground of artificial intelligence, as it is the perfect way to test the battle between human intuition and massive computing power. Due to the complex but well-defined nature of strategic games in general, they have proven to be a perfect environment for testing any progress in artificial intelligence.

Competing against other online players is a bigger part of chess competitions today than ever before, so is not surprising that online fair play is taken more seriously both by chess players who see cheating as a huge impediment to their entire game experience and by companies who create cheat detection software.

Cheat detection software has already been used for important games onboard, even before the COVID-19 pandemic, but this period of forced digitalization has meant that the implementation of cheat detection software has been done on a large scale, even for entertainment games.

We have reached a fundamental question about the entertainment value of chess: What do people really want to see—chess games played between chess players who are sometimes wrong (this being probably the most interesting and educational part to watch) or games played by chess engines, which are beautiful but painfully perfect? Moreover, how do we evaluate the entertainment value or the educational power of a chess game? It is important to analyze what beauty really is in chess. Algorithms for beauty detection in chess diagrams are currently being developed, but this is a topic that needs to be addressed in a future paper.

Conflicts of Interest

None declared.

References

1. Waldstein D. Chess thrives online despite pandemic. The New York Times. 2020 May 08. URL: https://www.nytimes.com/2020/05/08/sports/coronavirus-chess-online-tournament.html [accessed 2020-09-23]
2. Risi S, Preuss M. From Chess and Atari to StarCraft and Beyond: How Game AI is Driving the World of AI. Künstl Intell 2020 Feb 18;34(1):7-17. [doi: 10.1007/s13218-020-00647-w]
3. Hsu FH. Behind Deep Blue: Building the Computer that Defeated the World Chess Champion. Princeton, NJ: Princeton University Press; 2002.
4. Müller K, Schaeffer J. Man vs. Machine: Challenging Human Supremacy at Chess. Milford, CT: Russell Enterprises; 2018.
5. Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science 2018 Dec 06;362(6419):1140-1144. [doi: 10.1126/science.aar6404]
6. Sutton RS, Barto AG. Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press; 1998.
7. Sadler M, Regan N. Game Changer: AlphaZero's Groundbreaking Chess Strategies and the Promise of AI. Alkmaar, Netherlands: New In Chess; 2019.
8. Schrittwieser J, Antonoglou I, Hubert T, Simonyan K, Sifre L, Schmitt S, et al. Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model. arXiv. URL: https://arxiv.org/abs/1911.08265 [accessed 2020-11-30]
9. Thompson K. Retrograde Analysis of Certain Endgames. ICGA J 1986 Sep 01;9(3):131-139. [doi: 10.3233/icg-1986-9302]
10. Somers J. How the Artificial-Intelligence Program AlphaZero Mastered Its Games. The New Yorker. 2018 Dec 28. URL: https://www.newyorker.com/science/elements/how-the-artificial-intelligence-program-alphazero-mastered-its-games [accessed 2020-11-25]
11. Schormann C. Can online chess overcome cheating? ChessTech. 2020 May 24. URL: https://www.chesstech.org/2020/can-online-chess-overcome-cheating/ [accessed 2020-12-07]

Abbreviations

FIDE: International Chess Federation

XSL•FO

RenderX

XSL•FO
RenderX

XSL•FO

RenderX